THE AGGREGATION OF DIHYDRODIPICOLINATE SYNTHASE

A thesis in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

in the University of Canterbury

by Sophie Keziah Walker

2008

In memory of Sandra Jackson.

“R2-D2, you know better than to trust a strange computer.” – C3PO

ACKNOWLEDGEMENTS

Of all those people who have helped me during the course of my PhD, the biggest thanks goes to my supervisors; Jack Heinemann, Juliet Gerrard and Grant "I learn most of my science from Mythbusters" Pearce. Your support has been phenomenal. The lessons you have taught me, unparalleled.

A great deal of thanks is due to all those people whose equipment I have begged, borrowed or stolen. Particular thanks go to Geoff Jameson who so very gently broke the news to me that my Q90L crystals were in fact salt. Matt Perugini: Thank you for guiding me through the minefield that is analytical ultracentrifugation and for sharing your interpretation of "the sprinkler" to the sounds of Blondie. Thanks to the Croft Institute for facilitating discussion between members of Bio21 (Melbourne) and the Gerrard group.

To all past and present members of the purple bubble, thank you for your camaraderie. It is you who have made the last three years what they were. Particular thanks to Ren, Shaggy and Andy. Your advice and friendship have been invaluable. Dearest Jackie. I have never really understood why everyone calls you "the heat".

The competent members of staff of the School of Biological Sciences have been of great help during the course of this thesis. Thank you, Manfred, for your ability to spot fibrils at 20 yards. Thank you, Matt, for your wizardry, and Jason, for kindly tolerating a statistically challenged biochemist. Thank you, Gerrald, for being a kindred spirit and for the ride on your motorcycle. Thank you, Selwyn and Nicole, for always asking. Ashley: You are the reason I did this PhD. I owe you everything.

To all my friends, thank you for putting up with my inability to hold a conversation. And thank you for tolerating my incapacity for social interaction.

To Mum, Dad, Charlotte and Jemimah, thank you for your patience and interest. As recompense for all your support, I am not even going to ask you read this.

TABLE OF CONTENTS

ABBREVIATIONS

%	percent
~	approximately
≤	less than or equal to
≥	greater than or equal to
°C	degrees Celcius
A	alanine
Å	Ångstrom
A_{340}	absorbance at 340 nm
A_{600}	absorbance at 600 nm
AFM	atomic force microscopy
$Agg_{½}$	time it takes to reach half the maximum absorbance monitored at 340 nm (calculated from equation 8.1)
ALS	amyotropic lateral sclerosis
ANOVA	analysis of variance
amp^r	gene conferring ampicillin resistance
AU	absorbance units
AUC	analytical ultra-centrifugation
(*S*)-ASA	(*S*)-aspartate β-semialdehyde
$\beta\text{-}agg_{½}$	time it takes to reach half the maximum fluorescence monitored at 482 nm (calculated from equation 8.1)
bp	base pairs
BSA	bovine serum albumin
C	cysteine
$c(s)$	sedimentation coefficient
CD	circular dichroism
cDHDPS	DHDPS from *Corynebacterium glutamicum*
cDNA	circular DNA
cm	centimetre
cp	centipoise
cryo-EM	cryo electron microscopy
D	aspartic acid
Da	Daltons
DAP	*meso*-diaminopimelate
*dap*A	gene encoding dihydrodipicolinate synthase
*dap*B	gene encoding dihydrodipicolinate reductase
dH_2O	distilled/deionsed water

DHDP	dihydrodipicolinate
DHDPR	dihydrodipicolinate reductase
DHDPS	dihydrodipicolinate synthase
DMSO	dimethyl sulfoxide
DNA	deoxyribose nucleic acid
DSF	differential scanning fluorimetery
DTT	dithiothreitol
E. coli	*Escherichia coli*
ε	extinction coefficient
e.g.	for example
EDTA	ethylenediaminetetraacetic acid
EM	electron microscopy
ER	endoplasmic reticulum
f/f_0	frictional ratio
FPLC	fast performance liquid chromatography
g	gram
g	gravities
G	guanine
GFP	green fluorescent protein
H	histidine
HEPES	*N*-2-Hydroxyethylpiperazine-*N*′-2-ethane sulfonic acid
HIC	hydrophobic interaction chromatography
HTPA	(4*S*)-4-hydroxy-2,3,4,5-tetrahydro-(2*S*)-dipicolinate
I	isoleucine
IC_{50}	inhibitor concentration giving 50 % inhibition
IEX	ion exchange chromatography
IPTG	isopropyl β-D-thiogalactopyranoside
J	Joule
K	Kelvin
kan^r	gene conferring kanamycin resistance
kb	kilobase
k_{cat}	catalytic constant
k_{cat}^{app}	apparent catalytic constant
kDa	kiloDalton
kg	kilogram
K_m	Michaelis-Menten constant
K_m^{app}	apparent Michaelis-Menten constant
kV	kiloVolt
L	litre

L	leucine
lac	operon encoding β-galactosidase
LB	Luria-Bertani media
M	moles per litre
mA	milliAmp
mAU	milli-absorbance units
mg	milligram
min	minutes
µL	microlitre
mL	millilitre
mM	millimolar
µm	micrometre
mm	millimeter
MOPS	3-(*N*-morpholino)propanesulfonic acid
MRSA	methicillin resistant *Staphylococcus aureus*
mV	millivolts
η	buffer viscosity (poise)
NAD^+	nicotamide adenine dinucleotide
NADH	nicotamide adenine dinucleotide, reduced form
$NADP^+$	nicotamide adenine dinucleotide phosphate
NADPH	nicotamide adenine dinucleotide phosphate, reduced form
NAL	*N*-acetyl neuraminate lyase
ng	nanogram
nm	nanometre
NMR	nuclear magnetic resonance
NOG	*N*-Oxalylglycine
P	probability
P	Proline
ρ	density
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PDB	protein data bank
pH	$-\log_{10}$ of the hydrogen ion concentration
pI	isoelectric point
Q	glutamine
*rec*A	gene encoding the RecA protein
RFU	relative fluorescence units
rmsd	root mean square deviation
RNase A	ribonuclease A

rpm	revolutions per minute
S	Svedberg coefficient
SD	standard deviation
SDS-PAGE	sodium dodecyl sulfate polyacrylimide gel electrophoresis
sec	seconds
SEM	standard error of the mean
sHsps	small heat shock proteins
σ	sigma value
s_{min}	minimum for continuous sedimentation fitting
s_{max}	maximum for continuous sedimentation fitting
SOB	super-optimal broth
SSNMR	solid-state nuclear magnetic resonance
Sso AcP	α/β protein acylphosphataste from *Sulfolobus solfataricus*
T	threonine
$t_{½}$	half point of sigmoid curve (calculated from equation 8.1)
TAE	tris-acetyl EDTA electrophoresis buffer
TCA	trichloroacetic acid
TEM	transmission electron microscopy
tet^r	gene conferring tetracycline resistance
TEV	tobacco etch virus
TFE	2,2,2-Trifluoroethanol
ThT	thioflavin T
TIM	triosephosphate isomerase
T_m	transition unfolding temperature
Tris.HCl	tris(hydroxymethyl)methylamine
TTR	transthyretin
UV	ultraviolet
V	volts
V	valine
$\bar{v}$	partial specific volume
v/v	unit volume per unit volume
V_e	elution volume
W	tryptophan
w/v	unit weight per unit volume
X-gal	5-bromo-4-chloro-3-indol-β-δ-galactosidase
Y	tyrosine
ZYM5052	autoinduction media
Z_{agg}	aggregation score of Zyggregator

ABSTRACT

An increasing number of diseases are associated with protein misfolding, one type of which results in the formation of amyloid fibrils. This research has addressed the hypothesis that all proteins can form amyloid fibrils and investigates what factors protect proteins from forming these macromolecular assemblies. Most analyses of the aggregation propensity of proteins have been limited to the properties of the amino acid sequence, thus fail to consider the roles that higher levels of organisation play in protecting polypeptides from misfolding. The $(\alpha/\beta)_8$ barrels are a common class of proteins and have never been shown to form amyloid fibrils. This thesis aims to elucidate the characteristics that prevent $(\alpha/\beta)_8$ barrels from misfolding using *Escherichia coli* dihydrodipicolinate synthase (DHDPS), a homotetrameric $(\alpha/\beta)_8$ barrel protein, as a model. It is widely accepted that the precursor of amyloid fibrils is a partially folded species. It is hypothesised in this thesis that the $(\alpha/\beta)_8$ barrel fold protects the protein against this partial unfolding. This was tested by generating a catalogue of site-directed mutants of DHDPS and screening each of these in a range of pHs and ionic strengths. Amorphous aggregation propensity was assessed by monitoring light scattering at 340 nm and β-sheet specific aggregation was assessed using ThT. Thermal stability was monitored using DSF and CD spectroscopy. Crystallography was used to assess tertiary and quaternary structures and in the cases where crystal structures were not obtained, kinetics was used as a proxy indicator of correct folding and monomeric association. CD spectroscopy was also used to investigate the secondary structure of the DHDPS variants and analytical gel permeation liquid chromatography and AUC were used to confirm quaternary structure.

The stability and aggregation propensity of DHDPS and its variants were assessed under a range of pH and salt conditions. It was established through the characterisation of the wild-type protein that the predominant determinant of stability was, unsurprisingly, pH. This was a trend observed for all the variants described.

Affinity tags were used during the course of this research to facilitate and expedite the production of the protein variants. The introduction of tags containing a polyhistidine motif to DHDPS significantly altered some biophysical properties. Whilst the secondary and quaternary structures were found to be similar to the wild-type enzyme, the catalytic properties were changed. In addition to this, the propensity to aggregate was altered. The full-length polyhistidine tags increased the propensity of DHDPS to form β-sheet-specific aggregate, although this did not result in the formation of amyloid fibrils for most of the variants.

The Zyggregator algorithm was used to predict amino acid substitutions that would increase the aggregation propensity of DHDPS. It identified several amino acids, three of which were chosen for mutation and two of which were expressed in sufficient quantity for further study. DHDPS Q90L and A207V were characterised and the amino acid substitutions did not significantly alter the kinetic parameters of the enzyme. The crystal structure of A207V was solved and confirmed the results of the kinetic analysis demonstrating unchanged tertiary and quaternary structures. Both variants exhibited tertiary and quaternary structures similar to the wild-type enzyme, although Q90L contained more disorder than the wild-type enzyme. The thermal denaturation temperatures and aggregation propensities were also similar to wild-type, although the propensity for both variants to

form β-sheet-specific aggregates was reduced. The combinatorial effects of Q90L, A207V and the polyhistidine tags were assessed. This revealed that whilst most biophysical properties were unaffected, the β-sheet-specific aggregation propensity for pET M11 and pET 151/D-TOPO DHDPS Q90L and pET M11 DHDPS A207V, were significantly increased compared to the wild-type enzyme.

The evolutionary forces driving the association of the monomeric and dimeric subunits of DHDPS are undetermined. Investigation of two quaternary structure mutants (DHDPS Y107W and L197Y) revealed that the tetrameric nature of *E. coli* DHDPS is important for protein activity, stability and the prevention of aggregation. The combinatorial affects of the disrupted quaternary structure and the polyhistidine tags further increased the predisposition of DHDPS to form β-sheet-specific aggregates, resulting in the formation of linear aggregates with some characteristics of amyloid fibrils.

The additive affect of Q90L, Y107W and a polyhistidine tag was assessed and revealed that the major determinant in protein stability and prevention of amorphous and β-sheet specific aggregation is the quaternary structure.

This study demonstrates that the destabilisation of the quaternary structure of DHDPS can result in the formation of amyloid-like aggregates by an $(\alpha/\beta)_8$ barrel, the first example of an $(\alpha/\beta)_8$ barrel misfolding in such a way. This finding supports the assertion that all proteins can form amyloid fibrils.

CHAPTER 1

INTRODUCTION

It is widely believed that all proteins are able to form amyloid fibrils, a highly stable protein aggregate with an ordered β-sheet-rich structure. Despite this, one of the most common protein folds, the $(\alpha/\beta)_8$ barrel, has never been shown to do so. This research aimed to investigate whether $(\alpha/\beta)_8$ barrels could form amyloid fibrils and the factors potentially preventing them from doing so. This was achieved through changing residues hypothesised to be key to aggregate formation in *Escherichia coli* DHDPS, a model, homotetrameric, $(\alpha/\beta)_8$ barrel enzyme.

1.1 PROTEIN FOLDING AND MISFOLDING

The structure of proteins are critical to their function [1]. Following synthesis from the ribosome, the unfolded amino acid sequence must fold into the protein's native state. The folding process that occurs is implicitly encoded in the primary structure or sequence of amino acids in the polypeptide chain (U in figure 1.1), but often the cellular environment (*e.g.* temperature, pH and ionic strength) and a range of cellular processes such as the action of chaperone proteins [2] increases the reliability of the process. During the folding, the polypeptide chain can occupy a vast range of conformational states [1, 3]. The native state can be an unfolded peptide with biological function (*e.g.* D1-D4 fibronectin-binding protein from *Staphylococcus aureus* [4]) or a more compact, globular structure (N in figure 1.1). The latter often achieves its final conformation via the formation of partially folded intermediates [3, 5] (I in figure 1.1). Once in the native state, proteins can associate to form oligomeric structures and, in some cases, can form linear macromolecular assemblies with biological functionality (*e.g.* actin and microtubulin) [5]. The protein folding pathway cannot be considered a distinct series of steps, but a series of equilibria, of which protein misfolding is a part (see figure 1.1) [3, 6].

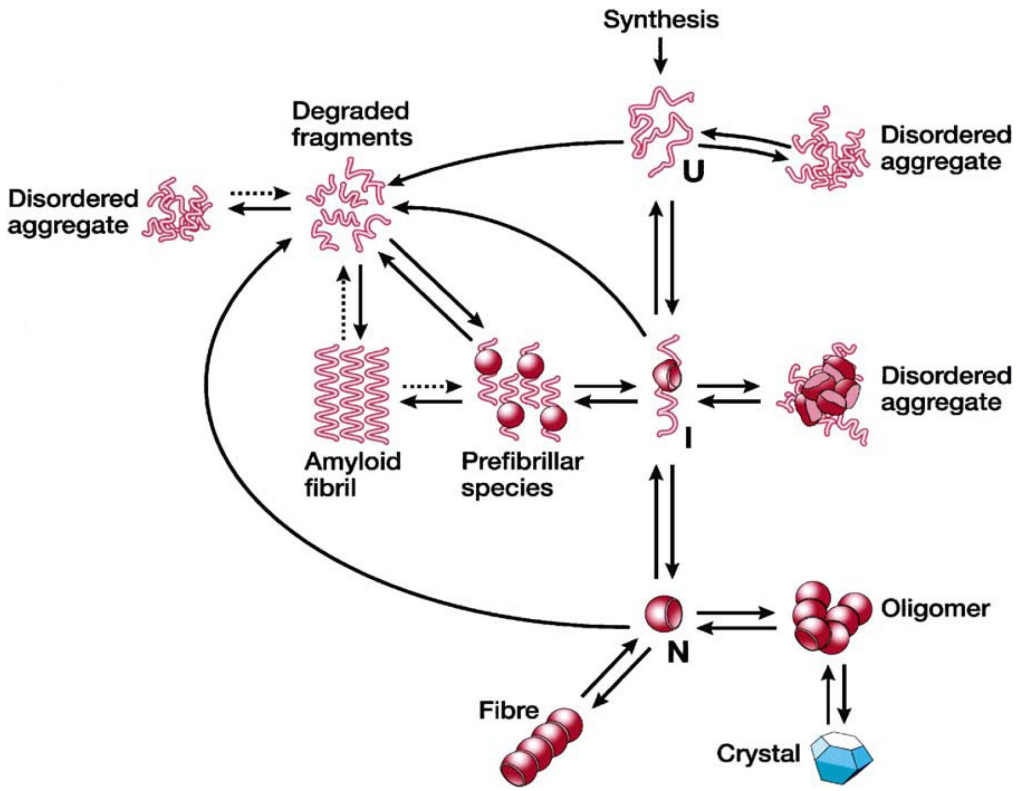


Figure 1.1 – A schematic representation of the folding pathway following synthesis of the polypeptide chain on the ribosome. Each species is in equilibrium with other conformations and various fates of the polypeptide chain are represented. U = the unfolded state, I = intermediate and N = native conformation. Figure taken from Dobson [3].

Protein folding events are entropically driven processes [7] and are commonly portrayed as a "landscape" of different conformations relative to their inherent stabilities (see figure 1.2) [6]. To reach an energetic minimum, the polypeptide has to follow the generally funnel-like landscape [7-9] to which intermediate structures confer "roughness". The number of amino acids in a peptide chain dictates the number of possible intermediate states and thus the roughness of the folding landscape [7, 10]. Small proteins are more likely to fold rapidly and correctly [11], while larger proteins have a wider range of potential intermediate states. Given the right circumstances, intermediate states can either fold into the native state or form alternate intermolecular interactions, entering into an aggregation funnel [12, 13] (see figure 1.2).

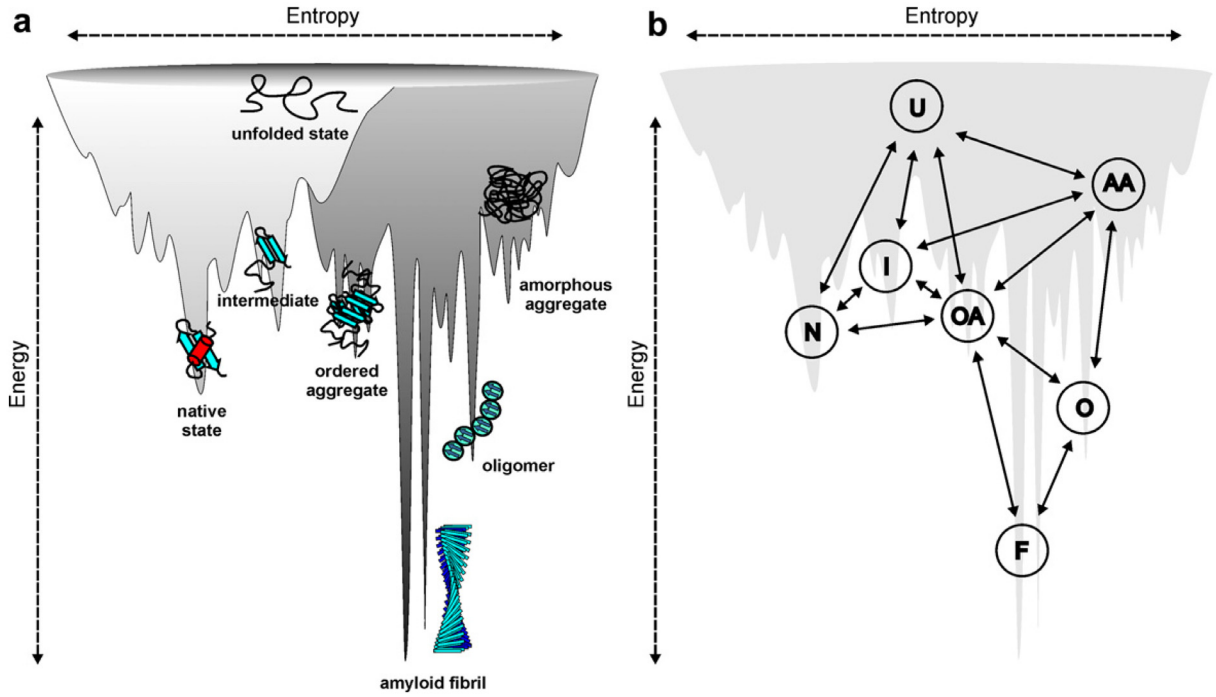


Figure 1.2 – **(a)** An example of a schematic energy landscape for protein folding including an aggregation funnel. The surface shows the funnelling of the various potential conformations of the polypeptide chain towards the native state (light grey), or through intermolecular interactions, the formation of native and non-native oligomers, amorphous aggregate, or amyloid fibrils (dark grey). Throughout the folding/misfolding process there are intermediate structures formed, some of which are represented in this diagram. **(b)** Proposed equilibria, linking various conformational states (U = unfolded, I = intermediate, N = native, OA = ordered aggregate, AA = amorphous aggregate, O = oligomer and F = fibril) in relation to their position on the energy landscape. Figure taken from Jahn and Radford [6].

The energetic minima encountered by the polypeptide during folding *in vivo* or *in vitro* are determined by the environment (*i.e.* the solution conditions, temperature and protein concentration) as well as by the sequence of the amino acid chain itself [6, 13, 14]. *In vivo,* folding processes are controlled by compartmentalisation (*e.g.* in the ER or mitochondria), the use of chaperones [1, 2, 15, 16], and evolutionary optimisation of proteins through enrichment of amino acid sequences with low aggregation propensities in their predominant physiological context [5, 17, 18]. These are all factors to ensure that a particular structure forms and its associated function is realised. However, aggregation events also represent energetic minima [13]. The relationship between the environment and the sequence also influences the propensity of proteins to enter into aggregation funnels,

resulting in either amorphous aggregation or the formation of more ordered aggregates called amyloid fibrils [6].

Amorphous aggregation occurs when the structure of proteins becomes perturbed and subunits misassemble. This affects both solubility and functionality of proteins and can sometimes, from a cell's point of view, have undesirable downstream affects. However, amorphous aggregates are usually able to be degraded or refolded both *in vivo* and *in vitro* through the action of proteases or chaperones [3]. Amyloid fibrils are far more stable structures and have been found to be resistant to most forms of degradation (*e.g.* heat, protease treatment), an attribute of the large number of hydrogen bonds in their highly ordered structure. It is these characteristics and the involvement of amyloid fibrils in normal and abnormal biological processes that have precipitated research into their properties.

1.2 AMYLOID FIBRILS IN BIOLOGY

An increasing number of diseases are associated with the misfolding of proteins [5, 6, 19-21]. Some of these disorders arise as a result of a reduction in the quantity of functional protein (*e.g.* cystic fibrosis) or aberrant protein trafficking [5]. However, the largest group of diseases associated with protein misfolding are the disorders involving aggregation. This aggregation can be amorphous in nature (*e.g.* in the case of α1-antitrypsin deficiency and amyotropic lateral sclerosis (ALS) [6]). Alternatively, the proteins or peptides can form fibrillar aggregates. These structures are known as amyloid fibrils and have been associated with several pathological conditions. These are classified as either neurodegenerative conditions (in which the aggregation occurs with in the brain), or "non-neuropathic amyloidoses" (in which the aggregation occurs in one or more tissues other than the brain) [5]. Amyloidoses can be grouped into three classes: sporadic, hereditary and transmissible forms. For example, Alzheimer's and Parkinson's disease can be either sporadic or hereditary. Huntingdon's disease, lysozyme and fibrinogen amyloidoses arise from inherited mutations, and some spongiform encephalopathies can be transmitted [5]. Amyloid fibrils were initially identified in the brain tissue of the patients suffering such neurological disorders and appeared as macroscopic abnormalities that exhibited positive iodine staining [22]. Recent advances have demonstrated that the primary constituent of amyloid fibrils is protein.

Although the association with pathologies is strong, there is now substantial evidence that amyloid-like fibrillar aggregates have biological functions in an increasing number of organisms. Examples of functional amyloid include: the use of curlin by *E. coli* to assist colonisation of inert surfaces, assisting biofilm formation and mediation of binding to host proteins [23], the use of hydrophobins by filamentous fungi to form degradation resistant, amphipathic monolayers on the surface of aerial structures [24], the formation of spider silk from spidronin by *Nephilla edulis* [25], the production of silk using chorion proteins by *Bombyx mori* (silkworm) [26], and the formation of fibrous striations inside human melanosomes by Pmel17 [27]. The number of proteins identified as forming fibrillar aggregates with a biological function is increasing and often these assemblies have the morphological, structural and dye-binding characteristics of amyloid fibrils. The range of organisms that functional amyloid has been observed in is diverse; however, one commonality is that the formation of these linear aggregates appears to be highly controlled [5].

Amyloid fibrils have also been shown to act as non-chromosomal genetic elements. Such proteins can exist in soluble conformations as well as in the form of a highly ordered linear aggregate with properties consistent with amyloid fibrils. As fibrils are a self-perpetuating structure, they can be inherited through non-Mendelian mechanisms. Proteins demonstrating such behaviour are called prions and although these are commonly associated with disease states in mammalian systems, they have been found to have a functional role in other organisms. For example, in *Saccharomyces cerevisae* the fibrillogenesis of the Sup35p protein confers the [PSI+] phenotype, facilitating the read-through of stop codons on the mRNA, resulting in the development of new phenotypes [28]. Another example is the Ure2p protein which, upon aggregation, prevents the binding and sequestering of the Gln3p transcription factor, resulting in the activation of a series of genes associated with nitrogen uptake in nitrogen limiting conditions. The strains of yeast exhibiting such aggregation are known as [URE3] [29]. A higher organism that uses amyloid fibrils as a non-chromosomal genetic element is the marine snail, *Aplysia californica,* in which a neuron-specific isoform of cytoplasmic polyadenylation element-binding (CPEB) protein which stimulates translation of CPEB-regulated mRNA, exhibits the highest activity in fibrillar form [30]. It has been hypothesised that this helps to maintain the long-term synaptic changes associated with memory storage [30].

Although it is their association with disease that has driven attempts to understand their structure and formation [6, 22, 31], the diverse biological roles of amyloid fibrils are increasingly recognised, and their potential use *in vitro* is starting to be explored [5, 32].

1.3 STRUCTURE OF AMYLOID FIBRILS

In order for aggregates to be classified as amyloid fibrils, they have to have certain biophysical properties. These properties include characteristic dye binding capacities and optical properties (*e.g.* birefringence in the presence of Congo red [33] and fluorescence in the presence of thioflavin T (ThT) [34-36]). Amyloid fibrils also exhibit a distinctive "cross-β" X-ray diffraction pattern. The reflections at 4.7 Å and 10 Å are generally thought to be a result of the β-sheet and inter-sheet packing [37-41]. Microscopic analysis of amyloid fibrils (by transmission electron microscopy (TEM) or atomic force microscopy (AFM)) have revealed that, typically, amyloid fibrils have a relatively linear morphology [39, 42] and are 60 – 300 Å in width [22, 42-44]. Analysis of the images has led to the hypothesis that fibrils consist of subunits called protofilaments [45]. The association of two or more protofilaments in a twisted conformation results in the formation of a mature fibril [42, 45-47] (see figure 1.3).

There have been several models proposed regarding the structure of amyloid fibrils [48, 49]. These include the hypotheses that the fibrils are formed from β-helices [50, 51] or water-filled nanotubes. This latter model has been refuted by the finding that hydrated and dehydrated fibrils of transthyretin (TTR) and lysozyme are similar, suggesting that the fibrils do not have a hollow core [52]. The prevailing model is that an amyloid fibril is comprised of a cross-β structure. According to this model, each protofilament consists of a "β-sheet helix" [37, 42, 53, 54] with β-strands running perpendicular to the axis of the fibril and with β-sheets running parallel to it [39, 40, 46] (see figure 1.3). These structures explain the characteristic X-ray diffraction patterns at 4.7 and 10 Å, respectively [37-41]. Recently, solid-state nuclear magnetic resonance (SSNMR) and X-ray crystallography have been used to investigate the secondary structure of amyloid fibrils. Generally, amyloid fibrils are not amenable to such techniques due to their large size, inherent insolubility and non-crystalline nature. However, short peptides with amyloid characteristics (*e.g.* a peptide derived from Sup35p, the amyloid β-peptide ($A\beta_{1-40}$ - the amyloid fibril forming region of the Aβ-protein which has been implicated in Alzheimer's disease) and the C-terminal domain of the fungal protein HET-s) [5] can be induced to crystallise, enabling the characterisation of the molecular packing of the cross-β spine by X-ray crystallography [40, 41]. Analysis of the fibrils formed by the Sup35p fragment has revealed that the two β-sheets formed by the peptide contribute to a single β-stand.

These β-stands are stacked in parallel, forming β-sheets that associate through the interaction of the amino acid side chains, excluding water, resulting in the formation of a "steric zipper", the composition and arrangement of which is thought to be dependent on the peptide sequence, orientation of the strands relative to one another, the side chains and the distance between the sheets [54-56]. What is clear from the models developed thus far is that, although amyloid fibrils can be formed from a diverse range of proteins (ranging from short peptides through to complete proteins), the mature fibrils all appear to have similar central arrangements [39]. They commonly have a cross-β structure and repetitive hydrophobic or polar interactions along the fibrillar axis and a core composed of between two and four closely associating sheets [57].

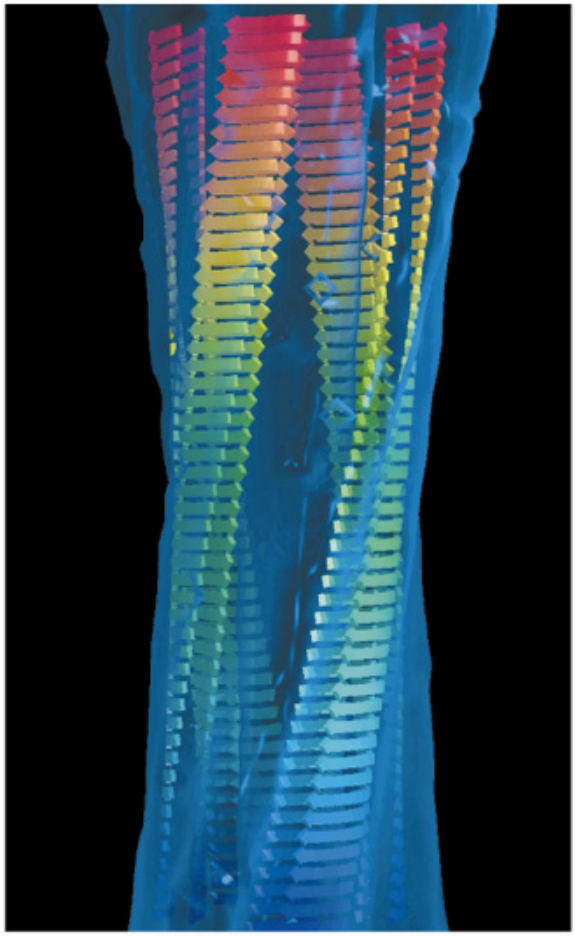


Figure 1.3 – The molecular model of an amyloid fibril developed from cryo-electron microscopy (cryo-EM) analysis of fibrils grown from an SH3 domain. This model considers the fibril to consist of four "protofilaments" that twist around each other to form a hollow tube with an internal diameter of ~ 60 Å [42]. This model represents a potential arrangement of regions of the polypeptide chain involved in the β-sheet structure assembly within amyloid fibrils (from Dobson [58]).

1.4 MECHANISMS OF AMYLOID FIBRIL FORMATION

In order to elucidate the structure and mechanisms underlying amyloid fibril formation, the conformational and oligomeric states adopted by a polypeptide chain during the misfolding pathway must be understood. The kinetic and thermodynamic processes that underlie the conformational changes required for the formation of all structures, including amyloid fibrils, are what dictate the outcome of the folding/misfolding pathway [59]. By identifying the intermediates involved in the formation of β-sheet-specific aggregates, the mechanisms underlying amyloid fibril formation can be further elucidated.

The most widely accepted model of amyloid fibril formation is the nucleated growth mechanism [5, 60]. This is characterised by a lag phase in which an aggregation nucleus is formed. This step is followed by rapid acceleration in fibril growth, eventually reaching a plateau [61]. The nucleus is a transient, thermodynamically unfavourable species, and its formation is considered the rate limiting step of fibril formation [5]. The formation of the nucleus is achieved through the presence of partially folded intermediates which, given the right environmental conditions, can associate to form the prefibrillar species (see figure 1.1). Upon nucleation, the fibrillation of the intermediates becomes more thermodynamically favourable, resulting in the exponential phase of the growth curve in which protofibrils are formed. These protofibrils associate to form the mature amyloid fibrils [31]. Support for the nucleated growth model arises from "cross seeding" experiments in which the lag phase can be shortened through the addition of pre-formed fibrillar species to a point where there is no discernable lag time. In such cases, the nucleation is not rate limiting [59, 62].

Some proteins appear to undergo amyloid formation in the absence of any lag phase and thus do not conform to the nucleated growth model. In such cases, it is possible that the rate limiting step is not the nucleation but another process in protein fibrillation [5]. It has been hypothesised that the formation of spherical oligomers and other prefibrillar forms can occur without a nucleation phase. Such misfolding events result in the formation of spherical particles or "worm-like" fibrils [57]. Other hypotheses consider oligomeric structures or off-pathway intermediates as precursors. In the case of β_2m microglobulin (associated with hemodialysis-related amyloidosis), pH plays a role in determining the nucleation times and thus fibril morphology. At pH 3.6, there appears to be no lag

time and the resulting fibrils exhibit a "worm-like" morphology while, at pH 2.5 the lag time is longer, consistent with a nucleated growth mechanism and long-straight fibrils are formed [57, 63].

Investigation into the formation of the prefibrillar species has revealed that amyloid formation by the Aβ-peptide proceeds *via* the formation of a structured protofilament. These have been visualised as spherical or beaded structures that bind Congo red and ThT and are distinct from the protofibrils that form the mature amyloid structure [5]. Another model of nucleation is that amyloid fibril formation is preceded by the formation of unstructured aggregates. For example the Sup35p protein from yeast has been shown to form oligomeric assemblies, containing significant β-sheet structures that are capable of nucleating amyloid fibril formation [64]. Similarly, amyloid formation can proceed *via* the partial unfolding of globular proteins [58, 59], a mechanistic model that arose from the observation that under destabilising conditions globular proteins were more likely to form amyloid fibrils than under non-destabilising conditions [65]. Conditions that can result in unfolding processes include changes in temperature and pH, high pressure, the presence of solvents, chaotropic agents and/or the presence of amino acid substitutions that destabilise the protein [5, 65-71].

Amyloid fibrils can also be formed from native-like oligomers (but, crucially, these are non-native). This model does not subscribe to the premise of the other models, which is that the destabilisation and resulting exposure of aggregation-prone regions results in intermolecular interactions [6] and is supported by the fact that several proteins have been shown to undergo amyloid formation without significant destabilisation of the folded protein [5]. An example of this is the α/β protein acylphosphataste from *Sulfolobus solfataricus* (Sso AcP) which undergoes fibrillation from a destabilised native-like state. The initial aggregates retain enzymatic activity and have few cross-β structures. The conversion of this precursor into a catalytically inactive form that binds ThT, Congo red and exhibits the characteristic cross-β diffraction pattern is relatively slow; however, it occurs more rapidly than the unfolding process. This suggests that, in this case, the formation of amyloid fibrils occurs from the reorganisation of the protein from a folded state [72]. Critically, it appears that the destabilisation of the native state is necessary for fibrillation to occur. Another mechanism of fibril formation involving aggregation from a native-like state is *via* domain swapping. An example of this is T7 endonuclease I which is a naturally domain-swapped dimeric protein which can form amyloid fibrils by a "runaway" domain swap [73].

1.5 INFLUENCE OF SEQUENCE ON AMYLOID FIBRIL FORMATION

Proteins are able to adopt a wide range of conformational states, consistent with their specific primary structure, which determines the behaviour of the protein under different environmental conditions. These conformations determine the behaviour of the protein in different environmental conditions, including the stability of the peptide and how prone the protein is to aggregation and amyloid formation [74, 75]. Polypeptides with different amino acid compositions exhibit different rates of amyloid formation, whether they initially exist as fully or partly folded species [5].

An important factor that determines the ability of polypeptide chains to aggregate is the hydrophobicity of the side chains of the amino acids. This is because hydrophobic residues tend to self associate and favour the transition of α-helices into β-sheets, facilitating aggregation and amyloid formation. Amino acid substitutions that introduce either more or fewer hydrophobic residues can increase or decrease the aggregation propensity of the protein [76-78]. In addition to this, evidence exists that there is evolutionary selection against both long sequences of hydrophobic residues and alternating patterns of hydrophilic and hydrophobic residues [17, 79]. The ability of amino acids to dictate the secondary structure of the peptide chain is thought to be important for influencing β-sheet-specific aggregation propensity [17, 56, 77, 80]. Patterns consisting of alternating hydrophilic and hydrophobic amino acid residues favour the formation of β-sheets which induce amyloid formation, thus are disfavoured by nature [17]. Another important characteristic of amino acids is the charge of the side chains. Highly charged groups are thought to hinder self-association processes [77, 81]. Chiti *et al.* [77] found that decreasing the net charge of a protein through amino acid replacement accelerated β-sheet-specific aggregation, and increasing the charge resulted in a species with a reduced propensity to form amyloid fibrils. Also the interaction of polypeptide chains with charged macromolecules can accelerate aggregation [56]. Unfolded proteins with biological functionality tend to be less hydrophobic and more charged than globular proteins, a characteristic thought to reduce the aggregation propensity of the protein under physiological conditions [82].

Single amino acid mutations can have significant effects on the aggregation of proteins [5] and have therefore been used to develop algorithms that predict the change in aggregation propensity upon amino acid substitution [83-85]. One such protein is human lysozyme, which has been implicated in amyloid-related disorders in humans. All reported cases have been associated with mutational

variants of lysozyme (*e.g.* I56T and D67H). These variants have been shown to be structurally similar to the wild-type protein; however, reduced thermal stability has been implicated in their elevated propensity to form amyloid fibrils. The formation of fibrils by lysozyme has been attributed to the presence of partially folded species which, under physiological conditions, are more likely to aggregate [86].

Building on these empirical observations, there are an increasing number of computational algorithms available (of which two are discussed in detail in section 1.6 and others in the conclusion). At the outset of this work, the most readily available algorithms were phenomenological models that predict aggregation rates of unfolded polypeptide chains based on physiochemical factors [87-89].

1.6 ALGORITHMIC PREDICTION OF AGGREGATION AND AMYLOID FORMATION

Recently, attempts have been made to predict protein β-sheet-specific aggregation and destabilisation with the view of improving the understanding of the mechanisms underlying protein aggregation and amyloid formation [74, 80, 85, 87, 90-100]. These predictions have resulted in the development of a number of computational algorithms [101-109]. The physiochemical properties of amino acids have been used to develop models that predict the absolute aggregation rates of the polypeptide as a whole and changes in aggregation propensity upon the substitution of residues (see section 1.5).

To understand and be able to predict protein aggregation and amyloid formation would provide clues for the treatment of the range of devastating neurological and metabolic disorders associated with protein misfolding [5, 19, 21, 55, 110]. Aggregation "hot spots" have already been identified [88]. These are regions of the amino acid chain that are predicted to be important for promoting amyloid formation. Upon mutation of residues within these regions, the β-sheet-specific aggregation can be influenced. These amino acid substitutions can be used to increase the amyloidogeneity of a protein in order to probe the mechanisms underlying misfolding [111, 112] or decrease the propensity of a protein to form amyloid fibrils, with the view of optimising biotechnological processes [113, 114].

The aggregation of unfolded polypeptide chains has been described [83, 89] and, based on this characterisation, the aggregation propensity of the protein as a whole can be calculated and the aggregation-prone regions identified [84, 88, 97, 107, 115]. Most of the current techniques employ phenomenological models, combining the physiochemical nature of the amino acid sequence with computational descriptions of β-sheet-specific aggregation. These models base their predictions on physiochemical processes [85, 87, 88] and experimental precedent [41, 116, 117], and consider that very short sections of peptide chains can act as catalysts for β-aggregation [118]. At the outset of this work, two algorithms designed to predict the overall aggregation propensity and aggregation "hot spots" were readily available and are described below. Other algorithms were published during the course of this work and are discussed in the concluding chapters of this thesis (chapter 7, section 7.2).

Zyggregator

The development of this algorithm began with quantitative analyses of the effects of mutations on the overall aggregation rates of polypeptides by Chiti *et al.* [83]. This revealed that the physicochemical characteristics of the amino acids could affect aggregation rates of polypeptides and that the propensity to form β-sheet-specific aggregate could be altered through changing specific residues. Further investigation into the roles of the physiochemical nature of the amino acids and their role in aggregation [83, 94] led to refinement of the predictions [88, 89]. Of particular importance in the calculations was the hydrophilic/hydrophobic patterning in a sequence which had been observed to favour the formation of β-sheet-rich structures [17]. The algorithm ultimately developed from the calculations was called Zyggregator (http://www-vendruscolo.ch.cam.ac.uk/zyggregator.php) and is designed to calculate the β-sheet-specific aggregation potential of polypeptides. It uses the characteristics of the amino acid sequence itself (*e.g.* charge, hydrophobicity, patterning of polar and non-polar residues and propensities of sequences to adopt secondary structural elements) and the extrinsic factors (*i.e.* ionic strength, pH and protein concentration) to predict the overall aggregation potential of the protein. It does this by comparing the protein sequence to 10,000 randomly generated sequences of the same length and calculating the comparative aggregation propensity. In addition to this, the algorithm can predict regions within the polypeptide sequence that are "sensitive" for aggregation. Single amino acid substitutions in these regions have been shown to dramatically alter aggregation rates (*e.g.* of peptides derived from Aβ, α-synuclein (involved in Parkinson's disease) and tau (associated with Alzheimer's)) resulting from the introduction or removal of repeating patterns [88, 119]. Thus, the algorithm also can be used to

identify amino acid replacements that increase or decrease the aggregation propensity of the protein. As most globular proteins that undergo amyloid fibril formation do so from an unfolded or partially folded state, the algorithm focuses on the aggregation propensity from these destabilised states [88, 89].

TANGO

TANGO (http://tango.crg.es/) is a statistical mechanics algorithm that predicts cross-β aggregation in denatured polypeptides and proteins. It considers each segment of the polypeptide and determines the likelihood of that segment forming β–turn, α–helix, β–sheet aggregation or α-helical aggregation and the accompanying desolvation penalty. From these factors and the physicochemical nature of the polypeptide chain and its environment, it identifies β-aggregating regions within the protein and predicts the effects of mutations on the aggregation propensity of the protein as a whole [87]. It achieves this through estimation of the percentage occupancy of β-aggregation conformation for each amino acid. If a polypeptide has a sequence of residues with at least 5 % β-aggregation occupancy, it is considered to have a tendency to form β-aggregates [120]. The difference between Zyggregator and TANGO is that TANGO makes an attempt to allow for native state stability. It achieves this through utilising the FOLD-X force field [120]; this is an algorithm concerned with the energetic parameters of proteins. It uses the thermodynamic properties of the predicted protein structure and the intrinsic flexibility or existence of an unfolded state to estimate the stability of proteins and protein complexes from the primary sequence [120]. Although there have been advances, the ability to accurately predict secondary and tertiary structures from primary sequence is still an elusive goal. Therefore, the accuracy of TANGO is limited by the precision of FOLD-X. Another difference between Zyggregator and TANGO is that TANGO assumes that the effect of patterning is negligible and it only provides relative aggregation propensities, thus it is not possible to calculate absolute aggregation rates using this algorithm.

The phenomenological models described above are cited as having a high degree of accuracy [106]. However, as the algorithms were tested and/or trained on the structures or sequences of proteins with well defined stabilities and known aggregation propensities, their accuracy with regard to proteins with un-described aggregation potentials must be further tested. Using *E. coli* DHDPS as a model protein (see section 1.10), the predictions of the algorithms were compared (see chapter 4, section 4.2). Due to collaboration between the Gerrard and Dobson groups, the algorithm

Zyggregator was readily available for use in routine analysis of protein aggregation propensities and prediction of amino acid mutations that would increase the amyloidogenicity of a model protein (DHDPS) and was thus used in this research.

1.7 INFLUENCE OF TERTIARY STRUCTURE ON AMYLOID FIBRIL FORMATION

Although primary structure (in combination with the physicochemical environment and the action of chaperones) ultimately dictates the structure of proteins and, therefore, their propensity to aggregate, it is challenging to use this information alone to predict the aggregation propensities of a protein that begins the aggregation process from its native fold. Since most proteins exist in a folded or partially-folded state [5], the tertiary and quaternary structure of proteins will inevitably influence their capacity for aggregation. For example, the native conformation of drk-SH3 (from *Drosophilla melanogaster*) exists in equilibrium with its unfolded state in non-denaturing conditions [121]. It is thought that the unfolded regions of proteins play critical roles in promoting aggregation, an assertion that is supported by the fact that for several proteins (*e.g.* Sup35p, Ure2p and HET-s), the regions responsible for forming the cross-β core of the amyloid fibril are present in the unstructured domain of the native protein [5, 122].

The tertiary structure often confers the biological functionality of proteins, whether they are partially unfolded or exist in a more compact, globular state. Reactive and hydrophobic groups are usually buried within the tertiary or quaternary structure of the protein; however, there is a balance between stability and functionality. For an enzyme to be catalytically active, functional groups need to be accessible. These exposed functional groups also increase the risk of non-specific interactions [123]. The hydrophobic residues on the surface of the protein are often involved in oligomerisation; however, if non-specific protein-protein interaction occurs, these same residues can facilitate aggregation.

Some specific classes of tertiary fold are over-represented in the Protein Data Bank (PDB), perhaps suggesting that the balance between the functionality and the stability are optimised. This leads to the hypothesis proposed in this thesis that certain structural families are more stable and less prone to

aggregation than others, whilst providing biologically functionality. However, the elucidation of tertiary structure is limited to proteins that form crystals or those that are small enough to be amenable to NMR, thus some families may be over-represented.

Although it has been asserted that all polypeptide chains (given the right conditions), are capable of misfolding and amyloid formation [124], not all classes of proteins have been shown to aggregate in such a way. Fibrils have been shown to form from a diverse range of proteins [125]. An example of a class of protein that, despite significant β-sheet content, has never been shown to form amyloid fibrils is the $(\alpha/\beta)_8$ barrels, a family of proteins that have been cited as the most common tertiary fold (and will discussed in section 1.9).

1.8 INFLUENCE OF QUATERNARY STRUCTURE ON AMYLOID FIBRIL FORMATION

In many proteins, sequences or structures with high aggregation propensities tend to be buried, thus avoiding exposure to environmental conditions that may precipitate the misfolding process [5]. This fact has led to the hypothesis that, in addition to providing catalytic function by minimising movement within the monomeric subunits [126, 127], the quaternary structure maintained by many macromolecular assemblies may assist in the prevention of unwanted aggregation and amyloid formation through burying hydrophobic residues in the protein-protein interface [16, 128]. As discussed above, there is a balance between forming a functional unit and the presence of hydrophobic residues on the exterior of the protein. The catalytic functionality of a protein may be improved by oligomerisation; however, the exposed hydrophobic residues necessary for such interactions may increase the aggregation propensity of the protein.

Transthyretin (TTR) is a protein in which the quaternary structure has been intimately linked with the prevention of amyloid fibril formation. TTR is a tetrameric protein which is implicated in both spontaneous and familial protein misfolding diseases in humans [129, 130]. The monomer is comprised of eight β-strands arranged into two β-sheets. Dimers form across a β-sheet interface between two monomers, and the tetramer forms upon the association of two dimers across a hydrophobic interface [131] (see figure 1.4).

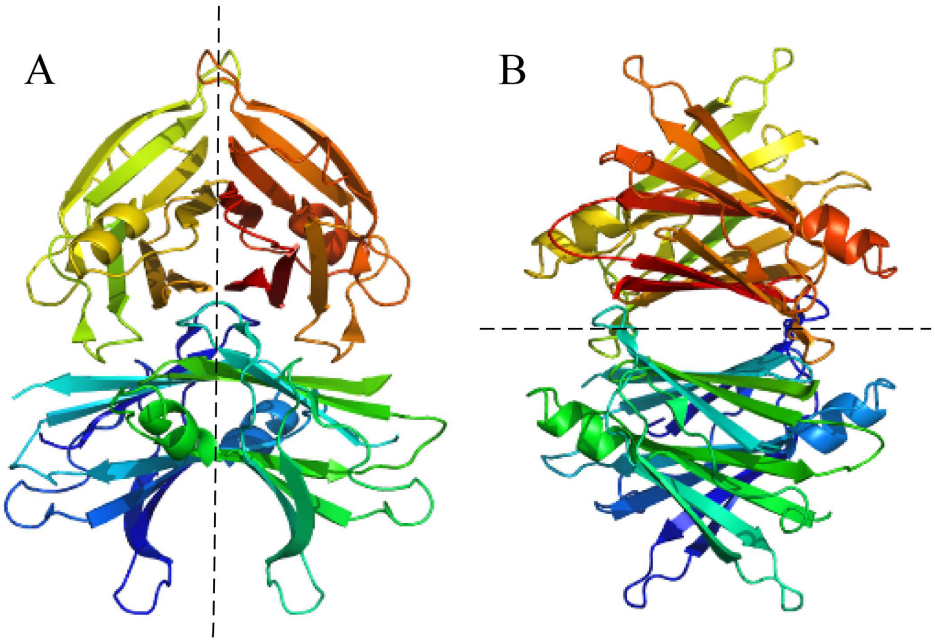


Figure 1.4 – TTR tetramer showing **(A)** the β-sheet interface and **(B)** the association of the dimers across the hydrophobic interface. This image was created from the TTR PDB file (accession number 1dvq) using PyMol [132].

The aggregation of TTR (associated with the clinical symptoms) is a consequence of the dissociation of the tetramer, upon which misfolding of the TTR monomers can occur, resulting in the formation of amyloid fibrils. In most cases, TTR amyloidoses are the result of single amino acid changes in the peptide chain. Several of these TTR variants have been described and the crystal structures have revealed that while some retain a structure similar to the wild-type enzyme, others have aberrant quaternary structures in the crystal form [133]. The destabilisation of the tetrameric state generally arises from the presence of amino acid substitutions in the loops involved in the hydrophobic interactions between the two dimers [129]. The dissociation of the dimers results in highly amyloidogenic TTR. Whether the monomeric or dimeric species is the precursor to fibrils has not been established. It has been hypothesised that the monomer is the precursor because, at low pH, TTR monomers have been shown to form fibrils at a rate > 100 times faster than the wild-type protein. However, the monomers exhibit a reduced tendency to misfold at neutral pH and partial denaturation of the monomer is required before amyloid fibril formation can occur. It has also been hypothesised that dimeric TTR is the precursor for amyloid fibril formation. Evidence supporting this includes the fact that dimers can form fibrils as well as cytotoxic spherical aggregates [131]. In addition to this, spin labelling experiments have shown that many of the interactions in the native

TTR dimer are present in TTR fibrils. It has been hypothesised that the extended β-sheet structure, present at the β-sheet interface between the monomeric subunits, acts as a catalyst for amyloid fibril formation [131]. The consensus between the models is that the quaternary structure of TTR protects it from forming amyloid fibrils [130]. The case study of TTR supports the hypothesis that, in the case of some proteins, the quaternary structure assists in the prevention of misfolding and amyloid fibril formation.

The stabilisation of monomers through buttressing against other subunits may lend some conformational stability in cases where the aggregation of globular proteins proceeds *via* the formation of native-like oligomers (*e.g.* Sso AcP, which only forms fibrils from a destabilised near-native state [72]). Well described inhibitors of aggregation (*e.g.* small heat shock proteins (sHsps) and extracellular clusterins) can also prevent aggregation through heterologous interactions, providing yet another example of protein-protein association conferring stability [16]. For example, sHsps act as the cell's first line of defence against protein aggregation brought about by physiological stress such as elevated temperature. They can prevent protein aggregation by binding to partially folded intermediates *via* hydrophobic interactions [128, 134]. This binding results in the formation of a chaperone-target complex which is both stable and soluble. sHsps cannot refold proteins; however, other chaperones can (*e.g.* Hsp70 and Hsp60), and these can act on the target proteins, refolding them into their native state. Alternatively, the stabilising capacity of sHsps allow the target protein to naturally refold, preventing aggregation [128, 135]. It is proposed here that interactions within oligomers act as a stabilising influence on the monomeric subunits of some proteins, thus demonstrating a "self-chaperoning" action.

1.9 $(\alpha/\beta)_8$ BARRELS

$(\alpha/\beta)_8$ Barrels consist of eight α-helices and eight β-sheets alternating along a peptide backbone. These structures are arranged in a "barrel-like" conformation or a "toroid". The β-sheets form the inner walls of the barrel while the α-helices form the outer wall. The interactions between the β-sheet core and the surrounding α-helices are hydrophobic in nature. The loops that link the α-helices and β-sheets generally comprise approximately 20 % of the peptide. The C-terminus generally contains the active site of the protein [136]. The size of $(\alpha/\beta)_8$ barrel proteins is variable with some existing as

small monomers (*e.g.* hevamine with 273 residues [137]), and some as large oligomeric structures (*e.g.* the tetrameric β-galactosidase, with 1023 amino acids per peptide chain [138]).

Despite being one of the most common protein folds, with as many as 10 % of the proteins in the PDB having an $(\alpha/\beta)_8$ barrel [139], the evolutionary advantage of this specific structure in terms of aggregation prevention is not clear. What is remarkable about this class of proteins is that there has not been a single report of any within the class forming amyloid fibrils. Thus the question arises, what factors prevent this class of proteins from misfolding? Here a model $(\alpha/\beta)_8$ barrel protein was used to investigate this question. This was the homotetrameric enzyme, DHDPS. This protein was chosen not only for its tertiary conformation, but also for the fact that through single amino acid substitution, the quaternary structure can be disrupted, resulting in both dimeric and monomeric/tetrameric species [126, 140], enabling investigation into the role of both the tertiary and quaternary structure in prevention of amyloid formation.

1.10 DIHYDRODIPICOLINATE SYNTHASE

DHDPS is the enzyme governing a branch point in the (*S*)-lysine biosynthetic pathway. It was initially identified in *E. coli* by Yugari and Gilvarg [141] and has subsequently been extensively characterised due to its involvement in metabolism and, more recently, due its potential as a target of novel antibiotics. It has been identified as a target because it is a key enzyme in the production of diaminopimelic acid (DAP) which is an important constituent of bacterial cell walls. For this reason, through inhibition of DHDPS, it has been hypothesised that bacterial growth may be inhibited [142].

DHDPS from bacteria is also of interest because lysine feedback inhibition is absent or occurs only at much higher concentrations of free lysine than is the case for plant enzymes. [143-152]. For this reason bacterial DHDPS has been used as a transgene in plants [153]. Lysine is often the limiting amino acid in animal diets, thus is often required as a nutritional supplement in the livestock industry. For decades there have been attempts to increase the lysine content of important food and feed crops through selective breeding [154]. More recently, corn has been genetically modified with a bacterial DHDPS sourced from *Corynebacterium glutamicum*. An example of this is the corn event LY038, which was engineered to express DHDPS from *C. glutamicum* (cDHDPS). Because DHDPS

is the enzyme involved in the rate-limiting step in the synthesis of lysine, the over-production of cDHDPS results in a corn line with elevated levels of free lysine and is therefore being proposed for use as an animal feedstock.

DHDPS is encoded by the *dap*A gene which has been identified in a range of plant and bacterial species [143, 151, 155-161]. The *dap*A genes from plant species are highly conserved at the DNA sequence level (~90 % sequence identity) whilst those from bacteria demonstrate some variability (~ 30 % sequence identity between some species). The tertiary structure is highly conserved across species. All exhibit very similar $(\alpha/\beta)_8$ barrel structures (see figure 1.5).

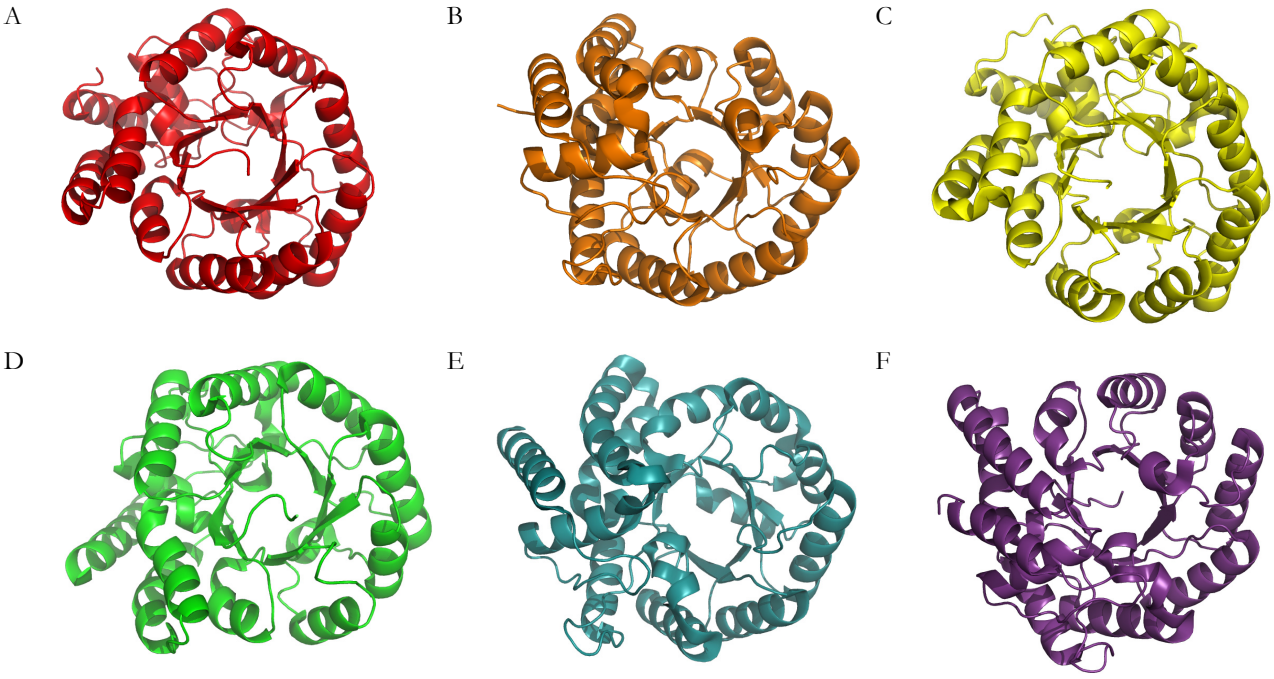


Figure 1.5 – The monomeric $(\alpha/\beta)_8$ barrel subunits of **(A)** *Thermotoga maritimea* (PDB accession code 1o5k), **(B)** *Mycobacterium tuberculosis* (PDB accession code 1XXX), **(C)** *Aquifex aeolicus* (PDB accession code 2ehh), **(D)** MRSA (PDB accession code 3daq), **(E)** *Oceanobacillus iheyensis* (PDB accession code 3d0c) and **(F)** *Bacillus anthrasis* (PDB accession code 1XKY). All the images were created using PyMol [132].

The crystal structure of *E. coli* DHDPS has been determined to a resolution of 1.9 Å and to 2.0 Å in a complex with lysine [162]. *E. coli* DHDPS exists as a homotetramer (see figure 1.6), comprised of four identical subunits each with a molecular mass of 31,272 Da [140, 162, 163]. The tetramer mass has been confirmed to be 125 kDa, as established though techniques such as gel filtration, blue

native PAGE, analytical ultracentrifugation (AUC), mass spectrometry and X-ray crystallography [140, 162-164]. Each monomer has two regions. The N-terminal domain is an $(\alpha/\beta)_8$ barrel, which is characterised by eight β-sheets surrounded by eight α-helices (see section 1.9 and figure 1.7). The C-terminal domain is a series of three α-helices [165, 166]. The catalytic site lies at the C-terminal end of the $(\alpha/\beta)_8$ barrel, a position typical for the $(\alpha/\beta)_8$ barrel enzymes [136].

The four $(\alpha/\beta)_8$ barrels are arranged as a dimer of dimers. Two of the monomers form a tight dimer interface which involves extensive hydrophobic interactions, while the loose dimer interface involves fewer residues (figure 1.6). Tetrameric *E. coli* DHDPS contains four active sites, one in each of the monomers [167]. The lysine binding site is positioned in the cleft between monomers responsible for the tight dimer interface. Two lysines are able to bind in each pocket and both form associations with both proximate monomers, resulting in feedback inhibition [162, 168, 169].

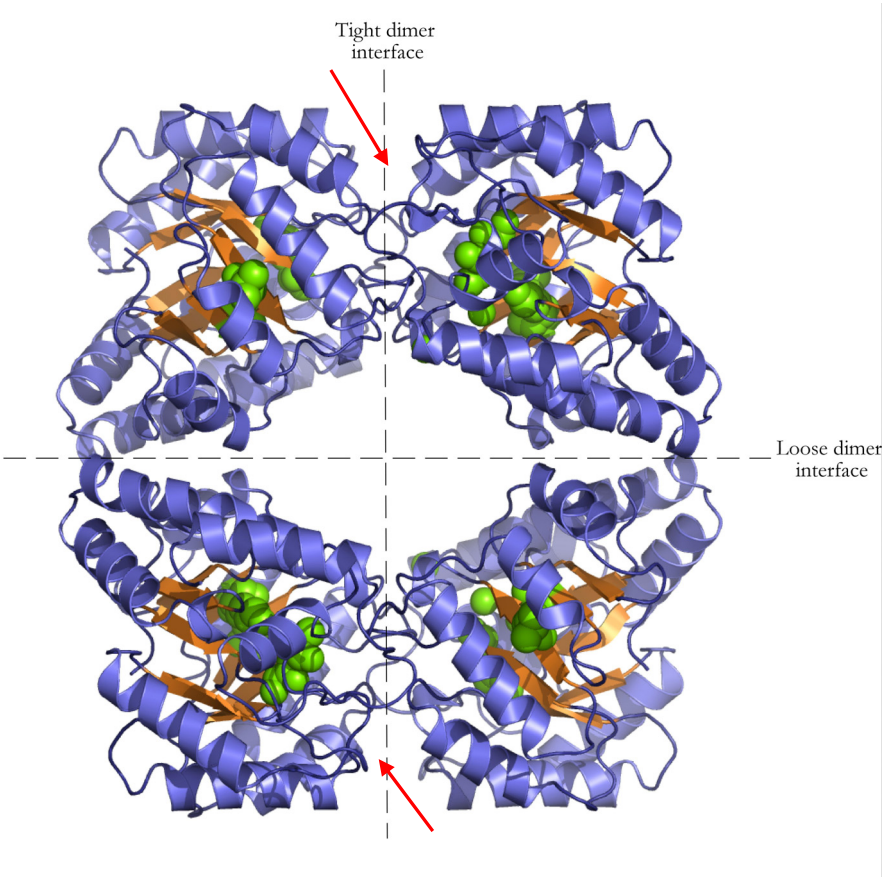


Figure 1.6 – Tetrameric DHDPS showing the α-helices (blue), β-sheets (orange) and residues involved in the active site (green). The red arrows indicate the position of the pockets in which the lysines bind. This image was generated from the PDB file (accession number 1yxc[162]) using PyMol [132].

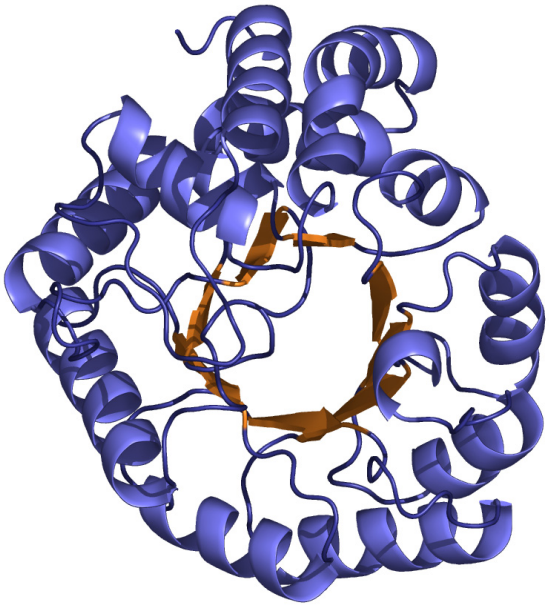


Figure 1.7 – The DHDPS $(\alpha/\beta)_8$ monomer with the α-helices (blue) and the β-sheets (orange) identified. This image was generated from the PDB file (accession number 1yxc [162]) using PyMol [132].

It appears that DHDPS uses a "ping-pong" mechanism of catalysis [170]. This requires the substrates to bind to and the products to be released from the enzyme in a defined order. DHDPS facilitates the condensation of (*S*)-aspartate-β-semialdehyde ((*S*)-ASA) and pyruvate, purportedly to form (4*S*)-4-hydroxy-2,3,4,5-tetrahydro-(2*S*)-dipicolinic acid (HTPA) (figure 1.8). The HTPA is then utilised by the enzyme dihydrodipicolinate reductase to form 2,3,4,5-tetrahydropicolinate with the conversion of NADPH to $NADP^+$ [171].

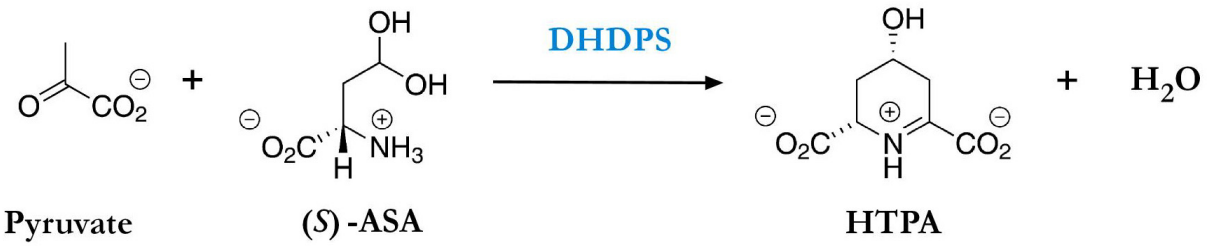


Figure 1.8 – The reaction catalysed by DHDPS (adapted from Dobson *et al.* [172]).

The quaternary structure of *Nicotiana sylvestris* DHDPS is like that of *E. coli* DHDPS because it consists of four monomers with a great deal of structural similarity to the *E. coli* enzyme. The difference between the *N. sylvestris* enzyme and *E. coli* DHDPS lies in the arrangement of the dimers. Whilst the tight dimer interfaces are similar, the loose dimer interface is reversed, with the opposing faces associating relative to *E. coli* DHDPS (figure 1.6) [173]. Upon disruption of the quaternary structure of DHDPS enzymes, the kinetic parameters and the mode of catalysis are altered [126, 140]. It was this observation that led to the hypothesis that the quaternary structure of DHDPS was essential for optimising catalysis by minimising the flexibility of the dimeric enzyme, achieved by the buttressing of the dimeric subunits in either the orientation observed for *E. coli* DHDPS or as observed for *N. sylvestris* DHDPS [126].

In all but two species characterised thus far, DHDPS is a tetramer. The exceptions are *Pisum sativum* which (according to preliminary gel filtration data), exists as a trimer [145] and methicillin-resistant (MRSA) and non-methicillin resistant *Staphylococcus aureus*, wherein DHDPS exist as dimers [174, 175]. In the MRSA dimer, the association of the monomeric subunits is driven by a larger number of contacts than is seen for *E. coli* DHDPS, potentially minimising the movement between the dimers, thus improving catalytic efficacy. The occurrence of the dimeric MRSA DHDPS provides evidence refuting the hypothesis that the tetramer is an essential evolutionary adaptation to the movement between the monomers and provides new impetus to investigate the evolutionary advantages of tetrameric DHDPS species. During the course of this thesis, the hypothesis that the tetrameric structure of *E. coli* DHDPS provides protection against aggregation and amyloid formation was investigated.

1.11 DO POLYHISTIDINE TAGS IMPACT ON AGGREGATION?

In order to determine the effects of the mutations on DHDPS, fairly large quantities of protein were required. Purification is often the greatest challenge in an experimental program designed to elucidate a protein's characteristics. Proteins are separated based on size, physicochemical properties, or binding affinities. The introduction of a polyhistidine tag onto a protein is a common method of altering the binding affinity of a protein to facilitate the purification of the protein of interest. A polyhistidine tag is an amino acid motif consisting of between 6 – 8 histidines and is either attached to the N- or C- terminus of a protein [176-180]. This is achieved either by inserting the gene of

interest into a vector that contains the tag, or by using primers containing the tag to attach the motif to the terminus of the protein [181, 182]. The polyhistidine tag enables the protein to bind to affinity media containing either nickel or cobalt ions [183, 184]. The tagged protein can be eluted from the media using a solution containing a high concentration of imidazole or histidine. The resulting protein is usually relatively pure after this single step procedure. In some cases, it is necessary to remove the polyhistidine tag prior to subsequent use [176, 177, 185]. This is achieved through the incorporation of a small number of amino acids between the protein itself and the polyhistidine tag that is recognised and cleaved by a specific endoprotease [176].

Polyhistidine tags are utilised widely in protein science because they facilitate protein purification. The vectors enabling the addition of polyhistidine tags are widely available commercially, thus making them readily available and effective tools for increasing and expediting protein production. Therefore these vectors were used during the course of this research to facilitate the production of *E. coli* DHDPS and its variants.

Since there is evidence that single amino acid substitutions can significantly alter the aggregation propensity of a protein as a whole (*e.g.* lysozyme and TTR exhibit increased tendencies to form amyloid fibrils as the result of single point mutations (discussed in sections 1.5 and 1.8)) [84, 88, 97, 107, 115] it seemed prudent to explore the impact of an additional sequence of amino acids. Despite the fact that polyhistidine tags are unstructured, have been shown to alter the stability of proteins [182], and contain hydrophobic sequences up to 10 amino acids in length, the effects of polyhistidine tags on amorphous aggregation and amyloidogeneity had never been the focus of a systematic investigation. As part of this research, the effect of polyhistidine tags on the propensity of DHDPS to form amorphous and β-sheet-specific aggregates was investigated.

1.12 OBJECTIVES

The aim of this research was to advance the understanding of the factors that prevent $(\alpha/\beta)_8$ barrel proteins from aggregating. This was achieved through the introduction of various point mutations into the sequence of *E. coli* DHDPS and the use of DHDPS variants that had already been shown to have a perturbed quaternary structure but had not been examined for their aggregation propensity.

It is thought that, for globular proteins, the precursor of amyloid fibrils is not the unfolded protein but a partially folded species. It is hypothesised that the $(\alpha/\beta)_8$ barrel provides protection against unfolding and subsequent aggregation. The role of this tertiary structure on the stability of DHDPS was investigated through the introduction of destabilising mutations, using Zygreggator to predict residues which played significant roles in determining the aggregation propensity of the protein. The biophysical characteristics of these point mutants was determined and compared with that of the wild-type.

Little is known with regard to the driving forces underlying the oligomerisation of proteins. In some cases the functional advantages are clear (*e.g.* haemoglobin [127]); however, the quaternary structure of DHDPS can be either a dimer (in the case of MRSA) or a tetramer (in the case of most other species). The evolutionary advantage of maintaining a tetrameric DHDPS is unclear. Through the introduction of amino acid substitutions that disrupt the oligomeric structure of DHDPS, the role of the quaternary structure preventing aggregation and amyloid fibril formation will be investigated.

1.13 CHAPTER SUMMARY

Chapter 2 introduces the techniques employed during the course of this thesis and presents the results of the biophysical characterisation of wild-type *E. coli* DHDPS. All the DHDPS variants were compared to the wild-type enzyme, thus the results presented in this chapter are used as "base-line" data.

Chapter 3 details the introduction of two different polyhistidine tags onto wild-type DHDPS and the changes in the biophysical properties caused by the extra amino acids associated with the motifs.

In chapter 4 the results of the algorithmic predictions are presented. These results were used to rationalise the creation of DHDPS Q90L, A207V and P270T which are characterised and compared to the wild-type enzyme. In addition, the combinatorial affects of Q90L, A207V and the polyhistidine tags were assessed.

Chapter 5 describes the results of the characterisation of DHDPS variants with disrupted quaternary structures arising from the Y107W and L197Y amino acid substitutions. In addition the combinatorial affects of the polyhistidine tags are assessed.

Chapter 6 describes the biophysical properties of polyhistidine-tagged DHDPS Q90L/Y107W. This double mutant combines both an amino acid substitution that disrupts the quaternary structure with one predicted to increase the β-sheet-specific aggregation of DHDPS.

Chapter 7 considers the results and conclusions that can be drawn from the preceding chapters.

Chapter 8 describes the materials and methods utilised during the course of this research.

1.14 REFERENCES

1 Dobson, C. M. (2003) Protein folding and misfolding. Nature **426**, 884 - 890

2 Barral, J. M., Broadley, S. A., Schaffar, G. and Hartl, F. U. (2004) Roles of molecular chaperones in protein misfolding diseases. Seminars in Cell and Developmental Biology **15**, 17 - 29

3 Dobson, C. M. (2004) Principles of protein folding, misfolding and aggregation. Seminars in Cell and Developmental Biology **15**, 3 - 16

4 Penkett, C. J., Redfield, C., Jones, J. A., Dodd, I., Hubbard, J., Smith, R. A. G., Smith, L. J. and Dobson, C. M. (1998) Structural and dynamical characterization of a biologically active unfolded fibronectin-binding protein from *Staphylococcus aureus*. Biochemistry **37**, 17054 - 17067

5 Chiti, F. and Dobson, C. M. (2006) Protein misfolding, functional amyloid and human disease. Annual Review of Biochemistry **75**, 333 - 366

6 Jahn, T. R. and Radford, S. (2008) Folding *versus* aggregation: Polypeptide conformations on competing pathways. Archives of Biochemistry and Biophysics. **469**, 100 - 117

7 Jahn, T. R. and Radford, S. (2005) The yin and yang of protein folding. FEBS Journal **272**, 5962 - 5970

8 Wolynes, P. G., Onuchic, J. N. and Thirumalai, D. (1995) Navigating the folding routes. Science **267**, 1619 - 1620

9 Dill, K. A. and Chan, H. S. (1997) From levinthal to pathways to funnels. Nature Structural Biology **4**, 10 - 19

10 Brockwell, D. J. and Radford, S. (2007) Intermediates: ubiquitous species on the folding energy landscape. Current Opinion in Structural Biology **17**, 30 - 37

11 Watters, A. L., Deka, P., Corrent, C., Callender, D., Varani, G., Sosnick, T. and Baker, D. (2007) Theory: The highly cooperative folding of small naturally occurring proteins is likely the result of natural selection. Cell **128**, 613 - 624

12 Vendruscolo, M. and Dobson, C. M. (2005) Towards complete descriptions of the free-energy landscapes of proteins. Philosophical Transactions of the Royal Society of London A **363**, 433 - 452

13 Clark, P. L. (2004) Protein folding in the cell: Reshaping the folding funnel. Trends in Biochemical Sciences **29**, 527 - 534

14 Gsponer, J. and Vendruscolo, M. (2006) Theoretical approaches to protein aggregation. Protein and Peptide Letters **13**, 287 - 293

15 Rousseau, F., Serrano, L. and Schymkowitz, J. (2006) How evolutionary pressure against protein aggregation shaped chaperone specificity. Journal of Molecular Biology **355**, 1037 - 1047

16 Carver, J. A., Rekas, A., Thorn, D. C. and Wilson, M. R. (2003) Small heat-shock proteins and clusterin: Intra- and extracellular molecular chaperones with a common mechanism of action and function. IUBMB Life **55**, 661 - 668

17 Broome, B. M. and Hecht, M. H. (2000) Nature disfavours sequences of alternating polar and non-polar amino acids: Implications for amyloidosis. Journal of Molecular Biology **296**, 961 - 968

18 Drummond, D. A., Bloom, J. D., Adami, C., Wilke, C. O. and Arnold, F. H. (2005) Why highly expressed proteins evolve slowly. Proceedings of the National Academy of Sciences USA **102**, 14338 - 14343

19 Stefani, M. (2004) Protein misfolding and aggregation: new examples in medicine and biology of the dark side of the protein world. Biochimica et Biophysica Acta **1739**, 5 - 25

20 Westermark, P., Benson, M. D., Buxbaum, J. N., Cohen, A. S., Frangione, B., Ikeda, S., Masters, C. L., Merlini, G., Saraiva, M. J. and Sipe, J. D. (2005) Amyloid: towards terminology clarification. Report from the Nomenclature Committee of the International Society of Amyloidosis. Amyloid **12**, 1 - 4

21 Stefani, M. and Dobson, C. M. (2003) Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution. Journal of Molecular Medicine **81**, 678 - 699

22 Sipe, J. D. and Cohen, A. S. (2000) Review: History of the amyloid fibril. Journal of Structural Biology **130**, 88 - 98

23 Chapman, M. R., Robinson, L. S., Pinker, J. S., Roth, R., Heuser, J., Hammer, M., Normark, S. and Hultgren, S. J. (2002) Role of *Escherichia coli* curli operons in directing amyloid fiber formation. Science **295**, 851 - 855

24 Mackay, J. P., Matthews, B. F., Winefield, R. D., MacKay, L. G., Haverkamp, R. G. and Templeton, M. D. (2001) The hydrophobin EAS is largely unstructured in solution and functions by forming amyloid-like structures. Structure **9**, 83 - 91

25 Kenney, J. M., Knight, D., J., W. M. and Vollrath, F. (2002) Amyloidogenic nature of spider silk. European Journal of Biochemistry **269**, 4159 - 4163

26 Iconomidou, V. A., Vriend, G. and Hamodrakas, S. J. (2000) Amyloids protect the silkworm oocyte and embryo. FEBS Letters **479**, 141 - 145

27 Fowler, D. M., Koulov, A. V., Alory-Jost, C., Marks, M. S., Balch, W. E. and Kelly, J. W. (2006) Functional amyloid within mammalian tissues. PLoS Biology **4**, 1 - 8

28 True, H. L. and Lindquist, S. L. (2000) A yeast prion provides a mechanism for genetic variation and phenotypic diversity. Nature **407**, 477 - 483

29 Chien, P., Weissman, J. S. and DePace, A. H. (2004) Emerging principles of conformation-based prion inheritance. Annual Review of Biochemistry **73**, 617 - 656

30 Si, K., Lindquist, S. L. and Kandel, E. R. (2003) A neuronal isoform of the *Aplysia* CPEB has prion-like properties. Cell **115**, 879 - 891

31 Pedersen, J. and Otzen, D. E. (2008) Amyloid - a state in many guises: Survival of the fittest fibril fold. Protein Science **17**, 2 - 10

32 Waterhouse, S. H. and Gerrard, J. A. (2004) Amyloid fibrils in bionanotechnology. Australian Journal of Chemistry **57**, 519 - 523

33 Frid, P., Anisimov, S. V. and Popovic, N. (2006) Congo red and the protein aggregation in neurodegenerative diseases. Brain Research Reviews **53**, 135 - 160

34 Groenning, M., Norrman, M., Flink, J. M., van der Weert, M., Butrinsky, J. T., Schluckebier, G. and Frokjaer, S. (2007) Binding of thioflavin T in insulin amyloid fibrils. Journal of Structural Biology **159**, 483 - 497

35 Eisert, R., Felau, L. and Brown, L. R. (2006) Methods for enhancing the accuracy and reproducibility of congo red and thioflavin T assays. Analytical Biochemistry **353**, 144 - 146

36 Voropai, E. S., Samtsov, M. P., Kaplevskii, K. N., Maskevich, A. A., Stepuro, V. I., Povarova, O. I., Kuznetsova, N., Turoverov, K. K., Fink, A. L. and Uversky, V. N. (2003) Spectral properties of thioflavin T and its complexes with amyloid fibrils. Journal of Applied Spectroscopy **70**, 868 - 874

37 Sunde, M., Serpell, L. C., Bartlam, M., Fraser, P. E., Pepys, M. B. and Blake, C. C. F. (1997) Common core structure of amyloid fibrils by synchrotron X-ray diffraction. Journal of Molecular Biology **273**, 729 - 739

38 Nelson, R. and Eisenberg, D. S. (2006) Recent atomic models of amyloid fibril structure. Current Opinion in Structural Biology **16**, 260 - 265

39 Makin, O. S. and Serpell, L. C. (2005) Structures for amyloid fibrils. FEBS Journal **272**, 5950 - 5961

40 Makin, O. S. and Serpell, L. C. (2005) X-ray diffraction studies of amyloid structure. Methods in Molecular Biology **299**, 67 - 80

41 Nelson, R., Sawaya, M. R., Balbirnie, M., Madsen, A. O., Riekel, C., Grothe, R. and Eisenberg, D. S. (2005) X-ray structure of amyloid cross-β spine. Nature **435**, 773 - 778

42 Jimenez, J. L., Guijarro, J. I., Orlova, E., Zurdo, J., Dobson, C. M., Sunde, M. and Saibil, H. R. (1999) Cryo-electron microscopy structure of an SH3 amyloid fibril and model of molecular packing. The EMBO Journal **18**, 815 - 821

43 Gosal, W. S., Myers, S. L., Radford, S. and Thomson, N. H. (2006) Amyloid under the atomic force microscope. Protein and Peptide Letters **13**, 261 - 270

44 Vestergaard, B., Groenning, M., Roessle, M., Kastrup, J. S., van de Weert, M., Flink, J. M., Frokjaer, S., Gajhede, M. and Svergun, D. I. (2007) A helical structural nucleus is the primary elongating unit of insulin amyloid fibrils. PLoS Biology **5**, 1089 - 1097

45 Makin, O. S. and Serpell, L. C. (2004) Structural characterisation of islet amyloid polypeptide fibrils. Journal of Molecular Biology **335**, 1279 - 1288

46 Serpell, L. C., Sunde, M., Benson, M. D., Tennent, G., Pepys, M. B. and Fraser, P. E. (2000) The protofilament substructure of amyloid fibrils. Journal of Molecular Biology **300**, 1033 - 1039

47 Jimenez, J. L., Nettleton, E. J., Bouchard, M., Robinson, C. V., Dobson, C. M. and Saibil, H. R. (2002) The protofilament structure of insulin amyloid fibrils. Proceedings of the National Academy of Sciences USA **99**, 9196 - 9201

48 Perutz, M. F., Finch, J. T., Berriman, J. and Lesk, A. (2002) Amyloid fibers are water-filled nanotubes. Proceedings of the National Academy of Sciences USA **99**, 5591 - 5595

49 Wille, H., Michelitsch, M. D., Guenebaut, V., Supattapone, S., Serben, A., Cohen, F. E., Agard, D. A. and Prusiner, S. B. (2002) Structural studies of the scrapie prion protein by electron crystallography. Proceedings of the National Academy of Sciences USA **99**, 3563 - 3568

50 Guo, J. T., Wetzel, R. and Xu, Y. (2004) Molecular modelling of the core of Aβ amyloid fibrils. Proteins **57**, 357 - 364

51 Tsai, H. H., Gunasekaran, K. and Nussinov, R. (2006) Sequence and structure analysis of parallel beta helices: implication for constructing amyloid structural models. Structure **14**, 1059 - 1072

52 Squires, A. M., Devlin, G. L., Gras, S. L., Ticker, A. K., MacPhee, C. E. and Dobson, C. M. (2006) X-ray scattering study of the effect of hydration on the cross-β structure of amyloid fibrils. Journal of the American Chemical Society **128**, 11738 - 11739

53 Sunde, M. and Blake, C. C. F. (1997) The structure of amyloid fibrils by electron microscopy and X-ray diffraction. Advances in Protein Chemistry **50**, 123 - 159

54 Sawaya, M. R., Sambashivan, S., Nelson, R., Ivanova, M. I., Sievers, S. A., Apostol, M. J., Thomson, M. J., Balbirnie, M., Wiltzius, J. J. W., McFarlane, H. T., Madsen, A. O., Riekel, C. and Eisenberg, D. S. (2007) Atomic structures of amyloid cross-β spine reveal steric zippers. Nature **447**, 453 - 457

55 Eisenberg, D. S., Nelson, R., Sawaya, M. R., Balbirnie, M., Sambashivan, S., Ivanova, M. I., Madsen, A. O. and Riekel, C. (2006) The structural biology of protein aggregation diseases: Fundamental questions and some answers. Accounts of Chemical Research **39**, 568 - 575

56 Fandrich, M. and Dobson, C. M. (2002) The behaviour of polyamino acids reveals an inverse side chain effect in amyloid structure formation. The EMBO Journal **21**, 5682 - 5690

57 Jahn, T. R., Parker, M. J., Homans, S. W. and Radford, S. (2006) Amyloid formation under physiological conditions proceeds via a native-like folding intermediate. Nature Structural and Molecular Biology **13**, 195 - 201

58 Dobson, C. M. (1999) Protein misfolding, evolution and disease. Trends in Biochemical Sciences **24**, 329 - 332

59 Uversky, V. N. and Fink, A. L. (2004) Conformational constraints for amyloid fibrillation: The importance of being unfolded. Biochimica et Biophysica Acta **1698**, 113 - 153

60 Xue, W.-F., Homans, S. W. and Radford, S. (2008) Systematic analysis of nucleation-dependent polymerization reveals new insights into the mechanism of amyloid self-assembly. Proceedings of the National Academy of Sciences USA **105**, 8926 - 8931

61 O'Nuallain, B., Shivaprasad, S., Kheterpal, I. and Wetzel, R. (2005) Thermodynamics of Aβ(1-40) amyloid fibril elongation. Biochemistry **44**, 12709 - 12718

62 Powers, E. T. and Powers, D. L. (2008) Mechanisms of protein fibril formation: Nucleated polymerization with competing off-pathway aggregation. Biophysical Journal **94**, 379 - 391

63 Smith, D. P., Jones, S., Serpell, L. C., Sunde, M. and Radford, S. (2003) A systematic investigation into the effect of protein destabilisation on β2-microglobulin amyloid formation. Journal of Molecular Biology **330**, 943 - 954

64 Serio, T. R., Cashikar, A. G., Kowal, A. S., Sawicki, G. J., Moslehi, J. J., Serpell, L. C., Arnsdorf, M. F. and Lindquist, S. L. (2000) Nucleated conformational conversion and the replication of conformational information by a prion determinant. Science **289**, 1317 - 1321

65 McParland, V. J., Kad, N. M., Kalverda, A. P., Brown, A., Kirwin-Jones, P., G., H. M., Sunde, M. and Radford, S. (2000) Partially unfolded states of β(2)-microglobulin and amyloid fibril formation *in vitro*. Biochemistry **39**, 8735 - 8746

66 Chiti, F., Bucciantini, M., Capanni, C., Taddei, F., Dobson, C. M. and Stefani, M. (2001) Solution conditions can promote formation of either amyloid protofilaments or mature fibrils from the HypF N-terminal domain. Protein Science **10**, 2541 - 2547

67 Guijarro, J. I., Sunde, M., Jones, L. J., Campbell, I. D. and Dobson, C. M. (1998) Amyloid fibril formation by an SH3 domain. Proceedings of the National Academy of Sciences USA **95**, 4244 - 4228

68 Silva, J. L., Foguel, D. and Royer, C. A. (2001) Pressure provides new insights into protein folding, dynamics and structure. Trends in Biochemical Sciences **26**, 612 - 618

69 Meersman, F., Dobson, C. M. and Heremans, K. (2006) Protein unfolding, amyloid fibril formation and configurational energy landscapes under high pressure conditions. Chemical Society Reviews **35**, 908 - 917

70 Hammarström, M., Jiang, X., Hurshman, A. R., Powers, E. T. and Kelly, J. W. (2002) Sequence-dependent denaturation energetics: A major determinant in amyloid disease diversity. Proceedings of the National Academy of Sciences USA **99**, 16427 - 16432

71 Chiti, F., Webster, P., Taddei, N., Clark, A., Stefani, M., Ramponi, G. and Dobson, C. M. (1999) Designing conditions for *in vitro* formation of amyloid protofilaments and fibrils. Proceedings of the National Academy of Sciences USA **96**, 3590 - 3594

72 Platoutski, G., Bemporad, F., Calamai, M., Taddei, F., Dobson, C. M. and Chiti, F. (2005) Evidence for a mechanism of amyloid formation involving molecular reorganisation within native-like precursor aggregates. Journal of Molecular Biology **351** 910 - 922

73 Guo, Z. and Eisenberg, D. S. (2006) Runaway domain swapping in amyloid-like fibrils of T7 endonuclease I. Proceedings of the National Academy of Sciences USA **103**, 8042 - 8047

74 de la Paz, M. L. and Serrano, L. (2004) Sequence determinants of amyloid fibril formation. Proceedings of the National Academy of Sciences USA **101**, 87 - 92

75 Gast, K., Modler, A. J., Damaschun, H., Krober, R., Lutsch, G., Zirwer, D., Golbik, R. and Damaschun, G. (2003) Effect of environmental conditions on aggregation and fibril formation of bastar. European Biophysical Journal **32**, 710 - 723

76 Wurth, C., Guimard, N. K. and Hecht, M. H. (2002) Mutations that reduce aggregation of the Alzheimer's Aβ42 peptide: an unbiased search for the sequence determinants of Aβ amyloidogenesis. Journal of Molecular Biology **319**, 1279 - 1290

77 Chiti, F., Calamai, M., Taddei, F., Stefani, M., Ramponi, G. and Dobson, C. M. (2002) Studies of the aggregation of mutant proteins *in vitro* provide insights into the genetics of amyloid diseases. Proceedings of the National Academy of Sciences USA **99**, 16419 - 16426

78 Otzen, D. E., Kristensen, O. and Oliveberg, M. (2000) Designed protein tetramer zipped together with a hydrophobic Alzheimer homology: A structural clue to amyloid assembly. Proceedings of the National Academy of Sciences USA **97**, 9907 - 9912

79 Schwartz, R., Istrail, S. and King, J. (2001) Frequencies of amino acid strings in globular protein sequences indicate suppression of blocks of consecutive hydrophobic residues. Protein Science **10**, 1023 - 1031

80 Kallberg, Y., Gustafsson, M., Persson, B., Thyberg, J. and Johansson, J. (2001) Prediction of amyloid fibril-forming proteins. The Journal of Biological Chemistry **276**, 12945 - 12950

81 Schmittschmitt, J. P. and Scholtz, J. M. (2003) The role of protein stability, solubility, and net charge in amyloid fibril formation. Protein Science **12**, 2374 - 2378

82 Uversky, V. N., Gillespie, J. R. and Fink, A. L. (2000) Why are "natively unfolded" proteins unstructured under physiologic conditions. Proteins: Structure, Function, and Genetics **41**, 415 - 427

83 Chiti, F., Stefani, M., Taddei, N., Ramponi, G. and Dobson, C. M. (2003) Rationalization of the effects of mutations on peptide and protein aggregation rates. Nature **424**, 805 - 808

84 Tartaglia, G. G., Cavalli, A., Pellarin, R. and Caflisch, A. (2004) The role of aromaticity, exposed surface, and dipole moment in determining protein aggregation rates. Protein Science **13**, 1939 - 1941

85 Tartaglia, G. G., Cavalli, A., Pellarin, R. and Caflisch, A. (2005) Prediction of aggregation rate and aggregation-prone segments in polypeptide sequences. Protein Science **14**, 2723 - 2734

86 Morozova-Roche, L. A., Zurdo, J., Spencer, A., Noppe, W., Receveur, V., Archer, D. B., Joniau, M. and Dobson, C. M. (2000) Amyloid fibril formation and seeding by wild-type

human lysozyme and its disease-related mutational variants. Journal of Structural Biology **130**, 339 - 351

87 Fernandez-Escamilla, A.-M., Rousseau, F., Schymkowitz, J. and Serrano, L. (2004) Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. Nature Biotechnology **22**, 1302-1306

88 Pawar, A. P., DuBay, K. F., Zurdo, J., Chiti, F., Vendruscolo, M. and Dobson, C. M. (2005) Prediction of "aggregation-prone" and "aggregation-susceptible" regions in proteins associated with neurodegenerative diseases. Journal of Molecular Biology **350**, 379-392

89 DuBay, K. F., Pawar, A. P., Chiti, F., Zurdo, J., Dobson, C. M. and Vendruscolo, M. (2004) Prediction of absolute aggregation rates of amyloidogenic polypeptide chains. Journal of Molecular Biology **341**, 1317 - 1326

90 Pastor, M. T., Esteras-Chopo, A. and de la Paz, M. L. (2005) Design of model systems for amyloid formation: Lessons for prediction and inhibition. Current Opinion in Structural Biology **15**, 57 - 63

91 de la Paz, M. L., Lacroix, E., Ramirez-Alvarado, M. and Serrano, L. (2001) Computer-aided design of β-sheet peptides. Journal of Molecular Biology **312**, 229 - 246

92 de la Paz, M. L., de Mori, G. M. S., Serrano, L. and Colombo, G. (2005) Sequence dependence of amyloid fibril formation: Insights from molecular dynamics simulations. Journal of Molecular Biology **349**, 583 - 596

93 Melquiond, A., Gelly, J.-C., Mousseau, N. and Derreumaux, P. (2007) Probing amyloid formation of the NFGAIL peptide by computer simulations. The Journal of Chemical Physics **126**, 065101 - 065107

94 Bemporad, F., Calloni, G., Camponi, S., Platoutski, G., Taddei, F. and Chiti, F. (2006) Sequence and structural determinants of amyloid fibril formation. Accounts of Chemical Research **39**, 620 - 627

95 Parthiban, V., Gromiha, M. M., Hoppe, C. and Schomburg, D. (2007) Structural analysis and prediction of protein mutant stability using distance and torsion potentials: Role of secondary structure and solvent accessibility. Protein: Structure, Function and Bioinformatics **66**, 41 - 52

96 Matysiak, S. and Clementi, C. (2006) Minimalist model as a diagnostic tool for misfolding and aggregation. Journal of Molecular Biology **363**, 297 - 308

97 Galzitskaya, O. V., Garbuzynskiy, S. O. and Lobanov, M. Y. (2006) Prediction of amyloidogenic and disordered regions in protein chains. PloS Computational Biology **2**, 1639 - 1648

98 Gilis, D. and Rooman, M. (1996) Stability changes upon mutation of solvent-accessible residues in proteins evaluated by database-derived potentials. Journal of Molecular Biology **257**, 1112 - 1126

99 Tartaglia, G. G., Pawar, A. P., Camponi, S., Dobson, C. M., Chiti, F. and Vendruscolo, M. (2008) Prediction of aggregation-prone regions in structured proteins. Journal of Molecular Biology **380**, 425 - 436

100 Yoon, S. and Welsh, W. J. (2004) Detecting hidden sequence propensity for amyloid fibril formation. Protein Science **13**, 2149 - 2160

101 Capriotti, E., Fariselli, P. and Casadio, R. (2005) I-Mutant2.0: Predicting stability changes upon mutation form the protein sequence or structure. Nucleic Acids Research **33**, W306 - W310

102 Gilis, D. and Rooman, M. (2000) PoPMuSIC, an algorithm for predicting protein mutant stability. Application to prion proteins. Protein Engineering **13**, 849 - 856

103 Parthiban, V., Gromiha, M. M. and Schomburg, D. (2006) CUPSAT: prediction of protein stability upon point mutation. Nucleic Acids Research **34**, W239 - W242

104 Kwasigroch, J. M., Gilis, D., Dehouck, Y. and Rooman, M. (2002) PoPMuSiC, rationally designing point mutations in protein structures. Bioinformatics **18**, 1701 - 1702

105 Stitziel, N. O., Binkowski, A., Tseng, Y. Y., Kasif, S. and Liang, J. (2004) topoSNP: a topographic database of non-synonymous single nucleotide polymorphisms with and without known disease association. Nucleic Acids Research **32**, D520 - D522

106 Caflisch, A. (2006) Computational models for the prediction of polypeptide aggregation propensity. Current Opinion in Structural Biology **10**, 437 - 444

107 Conchillo-Sole, O., de Groot, N. S., Aviles, F. X., Vendrell, J., Daura, X. and Ventura, S. (2007) AGGRESCAN: a server for the prediction and evaluation of "hot spots" of aggregation in polypeptides. BMC Bioinformatics **8**, 65

108 Zibaee, S., Makin, O. S., Goedert, M. and Serpell, L. C. (2007) A simple algorithm locates β-strands in the amyloid fibril core of alpha-synuclein, A beta, and tau using the amino acid sequence alone. Protein Science **16**, 906 - 918

109 Trovato, A., Seno, F. and Tosatto, S. C. E. (2007) The PASTA served for protein aggregation prediction. Protein Engineering, Design and Selection **20**, 521 - 523

110 Dobson, C. M. (2001) The structural basis of protein folding and its links with human disease. Philosophical Transactions of the Royal Society of London B **356**, 133 - 145

111 Fowler, S. B., Poon, S., Muff, R., Chiti, F., Dobson, C. M. and Zurdo, J. (2005) Rational design of aggregation-resistant bioactive peptides: reengineering human calcitonin. Proceedings of the National Academy of Sciences USA **102**, 10105 - 10110

112 Ventura, S. and Villaverde, A. (2006) Protein quality control in bacterial inclusion bodies. Trends in Biotechnology **24**, 179 - 185

113 Chalifour, R. J., McLaughlin, R. W., Lavoie, L., Morissette, C., Tremblay, N., Boule, M., Sarazin, P., Stea, D., Lacombe, D., Tremblay, P. and Gervais, F. (2003) Stereo selective interactions of peptide inhibitors with β-amyloid peptide. Journal of Biological Chemistry **278**, 34874 - 34881

114 Soto, C., Sigurdsson, E. M., Morelli, L., Kumar, R. A., Castano, E. M. and Frangione, B. (1998) β-sheet breaker peptides inhibit fibrillogenesis in rat brain model for amyloidosis: Implications for Alzheimer's disease. Nature Medicine **4**, 822 - 826

115 Thomson, M. J., Sievers, S. A., Karanicolas, J., Ivanova, M. I., Baker, D. and Eisenberg, D. S. (2006) The 3D profile method for identifying fibril-forming segments of proteins. Proceedings of the National Academy of Sciences USA **103**, 4074 - 4078

116 Williams, A. D., Shivaprasad, S. and Wetzel, R. (2006) Alanine scanning mutagenesis of Aβ(1-40) amyloid fibril stability. Journal of Molecular Biology **357**, 1283 - 1294

117 Jones, S., Manning, J., Kad, N. M. and Radford, S. (2003) Amyloid-forming peptides from β(2)-microglobulin. Insights into the mechanism of fibril formation *in vitro*. Journal of Molecular Biology **325**, 249 - 257

118 Esteras-Chopo, A., Serrano, L. and Lopez de la Paz, M. (2005) The amyloid stretch hypothesis: recruiting proteins towards the dark side. Proceedings of the National Academy of Sciences USA **102**, 16672 - 16677

119 Chiti, F., Taddei, F., Baroni, F., Capanni, C., Stefani, M., Ramponi, G. and Dobson, C. M. (2002) Kinetic partitioning of protein folding and aggregation. Nature Structural Biology **9**, 137 - 143

120 Guerois, R., Nielsen, J. E. and Serrano, L. (2002) Predicting changes in the stability of proteins and protein complexes: A study of more than 1000 mutations. Journal of Molecular Biology **320**, 369 - 387

121 Crowhurst, K. A., Tollinger, M. and Forman-Kay, J. D. (2002) Cooperative interactions and a non-native buried Trp in the unfolded state of an SH3 domain. Journal of Molecular Biology **322**, 163 - 178

122 Platt, G. W., Routledge, K. E., Homans, S. W. and Radford, S. (2008) Fibril growth kinetics reveal a region of β2-microglobulin important for nucleation and elongation of aggregation. Journal of Molecular Biology **378**, 251 - 263

123 Fersht, A. R. (2000) Structure and mechanism in protein science: A guide to enzyme analysis and protein folding. W. H. Freeman and Company, New York

124 Dobson, C. M., Sali, A. and Karplus, M. (1998) Protein folding: A perspective from theory and experiment. Angewandte Chemie International Edition **37**, 868 - 893

125 Aso, Y., Shiraki, K. and Takagi, M. (2007) Systematic analysis of aggregates from 38 kinds of non disease-related proteins: Identifying the intrinsic propensity of polypeptides to form amyloid fibrils. Bioscience, Biotechnology and Biochemistry **71**, 1313 - 1321

126 Griffin, M. D., Dobson, R. C. J., Pearce, F. G., Antonio, L., Whitten, A. E., Liew, C. K., Mackay, J. P., Trewhella, J., Jameson, G. B., Perugini, M. A. and Gerrard, J. A. (2008) Evolution of quaternary structure in a homotetrameric enzyme. Journal of Molecular Biology **380**, 691 - 703

127 Lukin, J. A. and Ho, C. (2004) The structure-function relationship of hemoglobin in solution at atomic resolution. Chemical Reviews **104**, 1219 - 1230

128 Rekas, A., Adda, C. G., Aquilina, J. A., Barnham, K. J., Sunde, M., Galatis, D., Williamson, N. A., Masters, C. L., Anders, R. F., Robinson, C. V., Cappai, R. and Carver, J. A. (2004) Interaction of the molecular chaperone αB-crystallin with α-synuclein: Effects of amyloid fibril formation and chaperone activity. Journal of Molecular Biology **340**, 1167 - 1183

129 Steward, R. E., Armen, R. S. and Daggett, V. (2008) Different disease-causing mutations in transthyretin trigger the same conformational conversion. Protein Engineering, Design and Selection **21**, 187 - 195

130 Hurschman Babbes, A. R., Powers, E. T. and Kelly, J. W. (2008) Quantification of the thermodynamically linked quaternary and tertiary structural stabilities of transthyretin and its disease-associated variants: The relationship between stability and amyloidosis. Biochemistry **47**, 6969 - 6984

131 Keetch, C. A., Bromley, E. H. C., McCammon, M. G., Wang, N., Christodoulou, J. and Robinson, C. V. (2005) L55P Transthyretin accelerates subunit exchange and leads to rapid formation of hybrid tetramers. Journal of Biological Chemistry **280**, 41667 - 41674

132 DeLano, W. L. (2002) The PyMOL molecular graphics system. DeLano Scientific, San Carlos

133 Damas, A. M. and Saraiva, M. J. (2000) Review: TTR amyoidosis - Structural features leading to protein aggregation and their implications on therapeutic strategies. Journal of Structural Biology **130**, 290 - 299

134 Ecroyd, H., Meehan, S., Horwitz, J., Aquilina, J. A., Benesch, J. L. P., Robinson, C. V., MacPhee, C. E. and Carver, J. A. (2007) Mimicking phyphorylation of αB-crystallin affects its chaperone activity. Biochemical Journal **401**

135 Hatters, D. M., Lindner, R. A., Carver, J. A. and Howlett, G. J. (2001) The molecular chaperone, α-crystallin, inhibits amyloid formation by apolipoprotein C-II. The Journal of Biological Chemistry **276**, 33755 - 33761

136 Wierenga, R. K. (2001) The TIM-barrel fold: a versatile framework for efficient enzymes. FEBS Letters **492**, 193 - 198

137 Terwisscha van Scheltinga, A. C., Hennig, M. and Dijkstra, B. W. (1996) The 1.8 Å resolution structure of hevamine, a plant chitinase/lysozyme, and analysis of the conserved sequence and structure motifs of glycosyl hydrolase family 18. Journal of Molecular Biology **262**, 243 - 257

138 Juers, D. H., Huber, R. E. and Matthews, B. W. (1999) Structural comparisons of TIM barrel proteins suggest functional and evolutionary relationships between β-galactosidase and other glycohydrolases. Protein Science **8**, 122 - 136

139 Copley, R. R. and Bork, P. (2000) Homology among $(\beta/\alpha)_8$ barrels: implications for the evolution of metabolic pathways. Journal of Molecular Biology **303**, 627 - 641

140 Pearce, F. G., Dobson, R. C., Weber, A., Lane, L. A., McCammon, M. G., Squire, M. A., Perugini, M. A., Jameson, G. B., Robinson, C. V. and Gerrard, J. A. (2008) Mutating the tight-dimer interface of dihydrodipicolinate synthase disrupts the enzyme quaternary structure: towards a monomeric enzyme. Biochemistry **In press**

141 Yugari, Y. and Gilvarg, C. (1965) The condensation step in diaminopimelate synthesis. The Journal of Biological Chemistry **240**, 4710 - 4716

142 Hutton, C. A., Perugini, M. A. and Gerrard, J. A. (2007) Inhibition of lysine biosynthesis: an evolving antibiotic strategy. Molecular BioSystems **3**, 458 - 465

143 Frisch, D. A., Gengenbach, B. G., Tommey, A. M., Sellner, J. M., Somers, D. A. and Myers, D. E. (1991) Isolation and characterization of dihydrodipicolinate synthase from maize. Plant Physiology **96**, 444 - 452

144 Ghislain, M., Frankard, V. and Jacobs, M. (1990) Dihydrodipicolinate synthase of *Nicotiana sylvestris*, a chloroplast-localized enzyme of the lysine pathmway. Planta **180**, 480 - 486

145 Dereppe, C., Bold, G., O, G., Ebert, E. and Schar, H. P. (1992) Purification and characterization of dihydrodipicolinate synthase from pea. Plant Physiology **98**, 813 - 821

146 Wallsgrove, R. M. and Mazelis, M. (1981) Spinach leaf dihydrodipicolinate synthase: Partial purification and characterization. Phytochemistry **20**, 2651 - 2655

147 Kumpaisal, R., Hashimoto, T. and Yamada, Y. (1987) Purification and characterization of dihydrodipicolinate synthase from wheat suspension cultures. Plant Physiology **85**, 145 - 151

148 Bartlett, A. T. M. and White, P. J. (1986) Regulation of the enzymes of lysine biosynthesis in *Bacillus sphaericus* NCTC 9602 during vegetative growth. Journal of General Microbiology **132**, 3169 - 3177

149 Karsten, W. E. (1997) Dihydrodipicolinate synthase for *Escherichia coli*: pH dependent changes in the kinetic mechanism and kinetic mechanism of allosteric inhibition by *L*-lysine. Biochemistry **36**, 1730 - 1739

150 Stahly, D. P. (1969) Dihydrodipicolinate acid synthase of *Bacillus licheniformis*. Biochimica et Biophysica Acta **191**, 439 - 451

151 Cremer, J., Eggeling, L. and Sahm, H. (1991) Control of the lysine biosynthesis sequence in *Corynebacterium glutamicum* as analyzed by overexpression of individual corresponding genes. Applied and Environmental Microbiology **57**, 1746 - 1752

152 Yamakura, F., Ikeda, Y., Kimura, K. and Sasakawa, T. (1974) Partial purification and some properties of pyruvate-aspartic semialdehyde condensing enzyme from sporulating *Bacillus subtilis*. Journal of Biochemistry **76**, 611 - 621

153 Dizigan, M. A., Kelly, R. A., Voyles, D. A., Luethy, M. H., Malvar, T. M. and Malloy, K. P. (2007) High lysine maize compositions and event LY038 maize plants. Monsanto Technology LLC, United States of America

154 Azevedo, R. A. and Lea, P. J. (2001) Lysine metabolism in higher plants. Amino Acids **20**, 261 - 279

155 Chen, N.-Y., Jiang, S.-Q., Klein, D. A. and Paulus, H. (1993) Organisation and nucleotide sequence of the *Bacillus subtilis* diaminopimelate operon, a cluster of genes encoding the first three enzymes of diaminopimelate synthesis and dipicolinate synthase. The Journal of Biological Chemistry **268**, 9448 - 9456

156 Garcia-Rodriguez, F. M., Zekri, S. and Toro, N. (2000) Characterization of the *Sinorhizobium meliloti* genes encoding a functional dihydrodipicolinate synthase (dapA) and dihydrodipicolinate reductase (dapB). Archives of Microbiology **173**, 438 - 444

157 Pisabarro, A., Malumbres, M., Mateos, L. M., Oguiza, J. A. and Martin, J. F. (1993) A cluster of three genes (dapA, orf2, and dapB) of *Brevibacterium lactofermentum* encodes

dihydrodipicolinate synthase, dihydrodipicolinate reductase, and a third polypeptide of unknown function. Journal of Bacteriology **175**, 2743 - 2749

158 Vauterin, M. and Jacobs, M. (1994) Isolation of a poplar and an *Arabidopsis thaliana* dihydrodipicolinate synthase cDNA clone. Plant Molecular Biology **25**, 545 - 550

159 Kaneko, T., Hashimoto, T., Kumpaisal, R. and Yamada, Y. (1990) Molecular cloning of wheat dihydrodipicolinate synthase. Journal of Biological Chemistry **265**, 17451 - 17455

160 Gunji, Y., Tsujimoto, N., Shimaoka, M., Ogawa-Miyata, Y., Sugimoto, S. and Yasueda, H. (2004) Characterization of the *L*-lysine biosynthetic pathway in the obligate methylotroph *Methylophilus methylotrophus*. Bioscience, Biotechnology and Biochemistry **68**, 1449 - 1460

161 Silk, G. W., Matthews, B. F., Somers, D. A. and Gengenbach, B. G. (1994) Cloning and expression of the soybean DapA gene encoding dihydrodipicolinate synthase. Plant Molecular Biology **26**, 989 - 993

162 Dobson, R. C. J., Griffin, M. D. W., Jameson, G. B. and Gerrard, J. A. (2005) The crystal structures of native and (*S*)-lysine-bound dihydrodipicolinate synthase from *Escherichia coli* with improved resolution show new features of biological significance. Acta Crystallographica Section D, Biological Crystallography **D61**, 1116 - 1124

163 Borthwick, E. B., Connell, S. J., Tudor, D. W., Robins, D. J., Shneier, A., Abell, C. and Coggins, J. R. (1995) *Escherichia coli* dihydrodipicolinate synthase: characterization of the imine intermediate and the product of bromopyruvate treatment by electrospray mass spectroscopy. Biochemical Journal **305**, 521 - 524

164 Griffin, M. D. W. (2005) PhD Thesis. Why is DHDPS a tetramer? University of Canterbury, Christchurch

165 Mirwaldt, C., Korndorfer, I. and Huber, R. (1995) The crystal structure of dihydrodipicolinate synthase from *Escherichia coli* at 2.5 Å resolution. Journal of Molecular Biology **246**, 227 - 239

166 Dobson, R. C. J., Gerrard, J. A. and Pearce, F. G. (2004) Dihydrodipicolinate synthase is not inhibited by its substrate, (*S*)-aspartate β-semialdehyde. Biochemical Journal **377**, 757 - 762

167 Blickling, S., Renner, C., Laber, B., Pohlenz, H.-D., Holak, T. A. and Huber, R. (1997) Reaction mechanism of *Escherichia coli* dihydrodipicolinate synthase investigated by X-ray Crystallography and NMR spectroscopy. Biochemistry **36**, 24 - 33

168 Coulter, C. V., Gerrard, J. A., Kraunsoe, J. A. E. and Pratt, A. J. (1999) *Escherichia coli* dihydrodipicolinate synthase and dihydrodipicolinate reductase: Kinetic and inhibition studies of two putative herbicide targets. Pesticide Science **55**, 887 - 895

169 Blickling, S. and Knablein, J. (1997) Feedback inhibition of dihydrodipicolinate synthase enzymes by L-lysine. Biological Chemistry **378**, 207 - 210

170 Gerrard, J. A. (1992) DPhil thesis. Studies on dihydrodipicolinate synthase, Oxford University, Oxford

171 Coulter, C. V. (1997) PhD thesis. Studies in lysine biosynthesis, University of Canterbury, Christchurch

172 Dobson, R. C. J., Griffin, M. D., Roberts, S. J. and Gerrard, J. A. (2004) Dihydrodipicolinate synthase (DHDPS) from *Escherichia coli* displays partial mixed inhibition with respect to its first substrate, pyruvate. Biochimie **86**, 311 - 315

173 Blickling, S., Beisel, H.-G., Bozic, D., Knablein, J., Laber, B. and Huber, R. (1997) Structure of dihydrodipicolinate synthase of *Nicotiana sylvestris* reveals novel quaternary structure. Journal of Molecular Biology **274**, 608 - 621

174 Burgess, B., Dobson, R. C. J., Bailey, M. F., Atkinson, S. C., Griffin, M. D. W., Jameson, G. B., Parker, M. W., Gerrard, J. A. and Perugini, M. A. (2008) Structure and evolution of a novel dimeric enzyme from a clinically-important bacterial pathogen. Journal of Biological Chemistry **283**, 27598 - 27603

175 Girish, T. S., Sharma, E. and Gopal, B. (2008) Structural and functional characterization of *Staphylococcus aureus* dihdrodipicolinate synthase. FEBS Letters **582**, 2923 - 2930

176 Arnau, J., Lauritzen, C., Petersen, G. E. and Pedersen, J. (2006) Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. Protein Expression and Purification **48**, 1 - 13

177 Kimple, M. E. and Sondek, J. (2004) Overview of affinity tags for protein purification. Current Protocols in Protein Science **Supplement 36**, 9.9.1 - 9.9.19

178 Sørensen, H. P. and Mortensen, K. K. (2005) Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. Journal of Biotechnology **115**, 113 - 128

179 Smith, M. C., Furman, T. C., Ingolia, T. D. and Pidgeon, C. (1988) Chelating peptide-immobilized metal ion affinity chromatography. A new concept in affinity chromatography. Journal of Biological Chemistry **263**, 7211 - 7215

180 Hochuli, E., Bannwarth, W., Dobeli, H., Gentz, R. and Stuber, D. (1988) Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent. Biotechnology **6**, 1321 - 1325

181 Cabrita, L. D., Dai, W. and Bottomley, S. P. (2006) A family of *E. coli* expression vectors for laboratory scale and high throughput soluble protein production. BMC Biotechnology **6**, 1 - 8

182 Dümmler, A., Lawrence, A.-M. and de Marco, A. (2005) Simplified screening for the detection of soluble fusion constructs expressed in *E. coli* using a modular set of vectors. Microbial Cell Factories **13**, 34

183 Porath, J. (1992) Immobilized metal ion affinity chromatography. Protein Expression and Purification **3**, 263 - 281

184 Porath, J., Carlsson, J., Olsson, I. and Belfrage, G. (1975) Metal chelate affinity chromatography, a new approach to protein fractionation. Nature **258**, 598 - 599

185 Kenig, M., Peternel, S., Gaberc-Porekar, V. and Menart, V. (2006) Influence of the protein oligomericity on final yield after affinity tag removal in purification of recombinant proteins. Journal of Chromatography A **1101**, 293 - 306

CHAPTER 2

CHARACTERISING AGGREGATION OF WILD-TYPE *ESCHERICHIA COLI* DHDPS.

2.1 INTRODUCTION TO METHODS

This chapter introduces the techniques employed during the course of this thesis and describes the biophysical properties of *E. coli* DHDPS. In subsequent chapters, the wild-type enzyme is compared to DHDPS variants using this suite of techniques. As discussed in chapter 1 section 1.1, the protein folding and misfolding pathways exist as a series of equilibria (figure 1.1). Each of the techniques described in this chapter provides information pertaining to one or more of these equilibria.

2.1.1 KINETICS

The progress of the reaction catalysed by DHDPS (section 1.10) was difficult to measure directly because neither the substrates nor the product of the condensation reaction which DHDPS catalyses were able to be detected spectrophotometrically. Thus the reaction of DHDPS was coupled to the next enzyme in the pathway, dihydrodipicolinate reductase (DHDPR), consuming NADPH, which could be monitored spectrophotometrically [1-7]. Since the rates of both reactions were coupled, the rate of the DHDPS-catalysed reaction could be inferred from the rate of the DHDPR-catalysed reaction in which HTPA is converted to dihydrodipicolinate (DHDP) (figure 2.1).

In order to measure the initial rate of DHDPS, DHDPR activity was kept in excess, ensuring that the reaction catalysed by DHDPS was the rate limiting step. As the oxidation of the NADPH to $NADP^+$ occurs at a 1:1 stoichiometry with the production of DHDP, the decrease in concentration of NADPH (monitored at 340 nm) could be used to calculate the activity of DHDPS.

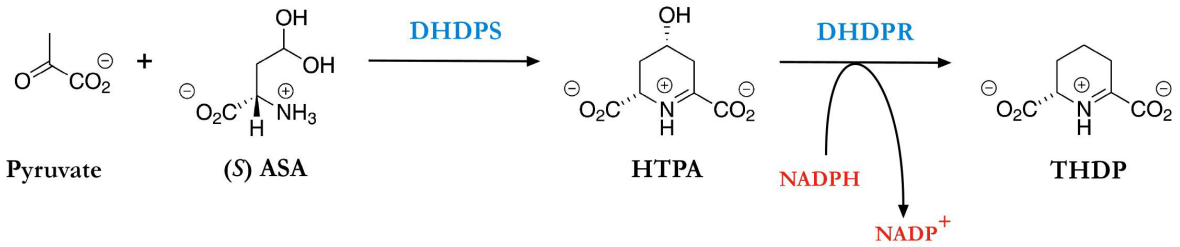


Figure 2.1 – The reactions catalysed by DHDPS and DHDPR. Adapted from Coulter *et al.*, 1999 [2].

The kinetic parameters of DHDPS have been shown to be dependent on both the tertiary structure and the tetrameric nature of the enzyme [8-10]. Disruption of the dimer-dimer interfaces disrupts the quaternary structure of DHDPS, attenuating the catalytic ability of these enzymes [8-10]. An unchanged kinetic profile, revealed by kinetic analysis, can be used as a proxy indicator for an enzyme with an intact active site and correct structure (figure 2.2).

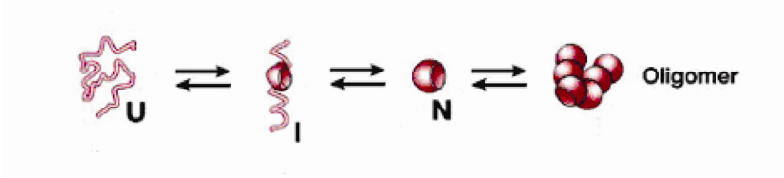


Figure 2.2 – The equilibria for which kinetic analysis provides information. Unchanged kinetic parameters for DHDPS indicate a correct tertiary and oligomeric structure. Modified from Dobson, 2004 [11]. U = the unfolded state, I = intermediate and N = the native conformation.

2.1.2 CRYSTALLOGRAPHIC ANALYSIS

The data obtained from the crystallographic analysis of a protein provides information relevant to all levels of the protein's physical state, from the primary sequence through to the quaternary structure. However, crystallographic data only provide information pertinent to the protein in crystal form (figure 2.3) therefore, complementary solution measurements are also required to understand the properties of the protein in solution.

Figure 2.3 – The equilibria involved in the formation of crystals, which provide information relevant to all levels of structure. Modified from Dobson, 2004 [11].

2.1.3 ANALYTICAL GEL PERMEATION LIQUID CHROMATOGRAPHY

Analytical gel permeation chromatography is a well established technique and can be used to study quaternary structure. It involves the approximation of macromolecular mass from the elution volume (V_e) [12]. The calculation of molecular mass provides basic information regarding the quaternary structure of proteins in solution (figure 2.2). Zonal analysis is the method generally used to estimate molecular mass.

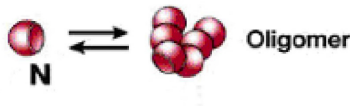


Figure 2.4 – The equilibria involved in the formation of oligomeric species. Modified from Dobson, 2004 [11].

2.1.4 ANALYTICAL ULTRA-CENTRIFUGATION

Analytical ultra-centrifugation (AUC) is a technique that has come back into vogue due to its ability to provide accurate information with regard to the size, mass, composition, interactions and distribution of macromolecular structures in solution [13, 14]. Unlike gel filtration, AUC does not involve the interaction of the protein with a matrix.

Two different measurements are possible through AUC. The first is sedimentation velocity. Applying large centrifugal forces to the solution and determining the position of the boundary formed as the

macromolecule concurrently experiences sedimentation and diffusion, reveals the hydrodynamic properties (*i.e.* size and shape) of the macromolecules in solution [13]. Sedimentation causes the protein to move to the distal end of the solution column over time [13]. Diffusion counteracts this, resulting in the formation of a boundary layer which can be identified using optical techniques. The data collected using this technique measures the migration of the boundary layer through the solution column. Both sedimentation and diffusion are dependent on the biophysical properties of the protein, such as the mass and the friction experienced by the protein in solution [13, 14]. It is therefore considered to be a more accurate method for the determination of the quaternary structures of proteins than analytical gel permeation liquid chromatography.

The second form of analysis is sedimentation equilibrium AUC which utilises lower centrifugal forces, enabling the balance of the sedimentation and diffusive forces, resulting in an apparently static boundary position. This technique provides more extensive information regarding the thermodynamic properties of the protein (*i.e.* molecular mass and equilibrium association constants for associating systems) [13-15]. As the equilibrium analyses are performed over a longer period of time (> 48 hours), a higher degree of protein stability is required than for velocity experiments. For this reason, only sedimentation velocity AUC analyses were performed in this thesis, as the proteins considered during the course of this study were designed to have an increased propensity to aggregate.

2.1.5 DIFFERENTIAL SCANNING FLUORIMETRY

Differential scanning fluorimetry (DSF) measures the fluorescence of a dye with emission properties that change upon interaction with unfolding proteins (figure 2.5) and has been used to identify ligands that bind and stabilise proteins, for drug screening and identification of conditions that promote stability of proteins [16-19]. During the course of this research, thermal denaturation assays were used to measure the thermal stability of proteins via direct measurement of the transition unfolding temperature (T_m). There are several dyes available for such assays [18, 20-22]. The dye SYPRO orange was used during the course of this research due to its desirable spectral properties (*i.e.* high quantum yields with low signal to noise ratio and stability of fluorescence emission in a range of environmental conditions) and high sensitivity (with the absence of any stabilising affect) [18, 20, 21]. SYPRO orange remains quenched, and thus non-fluorescent, in aqueous solution;

however, upon binding to hydrophobic regions of proteins, a large increase in fluorescence at 575 nm occurs resulting from excitation with light at 485 nm. Thus, as a protein unfolds and an increasing number of hydrophobic patches are exposed, the fluorescence increases [20, 21].

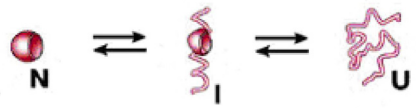


Figure 2.5 – DSF provides information regarding the thermal denaturation of proteins in solution. Modified from Dobson, 2004 [11].

2.1.6 CIRCULAR DICHROISM SPECTROSCOPY

Circular dichroism (CD) spectroscopy provides information regarding the secondary structure of proteins (figure 2.6) as well as thermal stability (figure 2.5) [23-25]. CD spectroscopy is a technique that measures differential absorption of left and right hand polarised light. The differential absorption is a result of structural asymmetry present in a solution of macromolecules through which the polarised light is passed. A structure containing no asymmetry has no CD intensity. Structures with asymmetry can result in both positive and negative CD intensities, depending on the structures present. The secondary structure of proteins can be assessed by monitoring wavelengths between 190 – 250 nm. α-Helix, β-sheet and random coil structures each give rise to distinct CD spectra, which allows the approximation of the proportion of each secondary structure present in a protein [23, 26, 27].

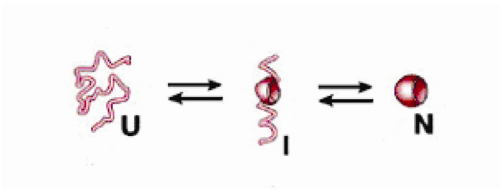


Figure 2.6 – CD wavelength scans provide information relevant to how the protein is folded in solution. Modified from Dobson, 2004 [11].

CD spectroscopy can also be used to investigate the thermal denaturation of proteins (figure 2.5). The unfolding of the proteins can be monitored by following changes in the CD spectra with

increasing temperature. A specific wavelength can be monitored, corresponding to a change in the population of a specific secondary structure [24].

2.1.7 AMORPHOUS AGGREGATION

Amorphous aggregation can be quantified by monitoring the absorbance at 340 nm (A_{340}). An increase in absorbance is indicative of increased light scattering due to the formation of protein precipitate [28-30] (figure 2.7). This technique was chosen as it has been shown to be useful in studies involving both disordered and β-sheet-specific aggregation [31], providing information regarding the formation of insoluble structures in solution.

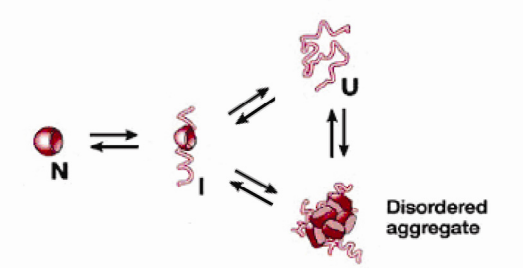


Figure 2.7 – The A_{340} assay provides information regarding the formation of amorphous aggregate in solution. Modified from Dobson, 2004 [11].

2.1.8 β-SHEET-SPECIFIC AGGREGATION

Thioflavin T (ThT) is a dye used to monitor the formation of β-sheet-specific aggregate as it fluoresces upon binding to protein structures with high β-sheet content [32-36]. Binding of ThT to proteinaceous β-sheet structures results in increased fluorescence at 482 nm upon excitation at 450 nm. An increase in fluorescence is considered indicative of the presence of amyloid fibrils (figure 2.8) although non-specific binding and fluorescence has been reported in some systems, particularly for proteins with naturally high β-sheet content [37, 38].

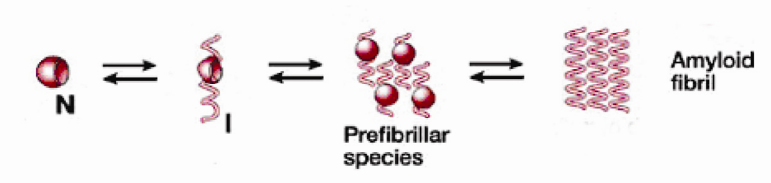


Figure 2.8 – The ThT assay provides information regarding the formation of β-sheet-specific aggregate in solution. Adapted from Dobson, 2004 [11].

2.1.9 CONFIRMATION OF AMYLOID FIBRILS

As ThT is not an absolute indicator of the presence of amyloid fibrils [33, 37, 38], confirmation of the presence of fibrillar structures must be confirmed through further biophysical techniques such as transmission electron microscopy (TEM) and X-ray fibre diffraction [38-41]. The examination of samples by TEM can provide microscopic confirmation of the presence of fibrillar structures. The confirmation of the presence of fibrils by X-ray fibre diffraction involves the construction of a "stalk" (chapter 8, section 8.8) and examination of the diffraction pattern obtained upon exposure to an X-ray beam. The presence of amyloid fibrils is confirmed by the appearance of two diffraction patterns at approximately 4.7 and 9.6 Å, characteristic of the cross-β pattern of the structure [42, 43].

2.2 SCREENING CONDITIONS

A standard series of conditions was used to assess the effect of temperature, pH and salt concentration on the results of the thermal stability, amorphous and β-sheet-specific aggregation assays performed on DHDPS. Through comparative analysis of the results of these assays, the relationship between stability, amorphous and β-sheet-specific aggregation of wild-type enzyme has been investigated. The results also provide baseline data to probe the affects of mutations on the misfolding of DHDPS (described in later chapters).

The assays were performed at pH 2, pH 4.5, pH 7, pH 9.5 and pH 11. These conditions were chosen because fibrils have been shown to form at a wide range of pH values [44-46]. Some proteins have been shown to form β-sheet rich aggregate at pH values close to their isoelectric point [44, 46, 47] whereas others are more inclined to form amyloid fibrils at pH values above or below their

isoelectric point [44, 45, 47]. As DHDPS has never been shown to form amyloid fibrils, a wide range of pH values was screened. In addition to the effect of net charge on the folded protein, pH is involved in the folding and unfolding of proteins [48, 49] as well as determining the formation of amorphous aggregates [44, 45, 50, 51]. The presence of unfolded species and/or amorphous aggregates is considered a key process in the formation of amyloid fibrils [44, 51-53]. The exposure of DHDPS to a wide range of pH conditions ensured the presence of various states of unfolding and misfolding in the assays.

Salt concentration can affect the solubility, stability and structure of proteins [54-63]. At low concentrations, salts typically stabilise proteins by non-specific electrostatic interactions (salting-in). At high concentrations, salt can either stabilise or destabilise proteins, depending on the salt and its concentration. Usually the abundance of salt ions decreases their solvating power observed at lower concentrations, reducing the solubility of the proteins [55]. Salt has also been shown to play a significant role in both promoting and preventing the formation of amorphous aggregates and amyloid fibrils. This is thought to be brought about by critical anion interaction leading to the necessary electrostatic and hydrophobic balance critical for amyloid formation [50, 64]. To maximise the likelihood of discovering this balance, thus promoting amyloid fibril formation of DHDPS, a range of salt conditions (0 mM NaCl, 20 mM NaCl, 40 mM NaCl, 100 mM NaCl and 500 mM NaCl) was screened. This ensured the likelihood of the formation of a series of destabilised and misfolded species [55, 64]. As salt and pH appear to have a combinatorial effect on the formation of fibrils [47, 64], all permutations of pH and salt concentrations were examined. Where possible, the influences of salt and pH on the equilibria (figure 2.1) were investigated using each method; however, experimental pragmatism did limit the numbers of conditions screened for some techniques. In such cases, a subset of conditions was chosen encompassing as many pH and salt concentrations as feasibly possible.

2.3 RESULTS

2.3.1 PROTEIN PURIFICATION

The purification of wild-type DHDPS was carried out according to the method adapted from Dobson, 2004 [65] as detailed in chapter 8, sections 8.4.1, 8.4.4 – 8.4.7. Following purification by ion

exchange chromatography (IEX) and hydrophobic interaction chromatography (HIC) (the results of which are shown by SDS-PAGE in figure 2.9), the average yield was 19.6 mg of DHDPS (76 units with a specific activity of 3.83 units/mg) per litre of bacterial culture, which is very similar to the yields and activity obtained by others [8, 66]. Multiple purifications were performed as it was observed that the stability of DHDPS was unreliable after approximately a 6 month period at -20 ºC [67, 68]. Where possible, all assays were carried out on protein obtained from a single purification.

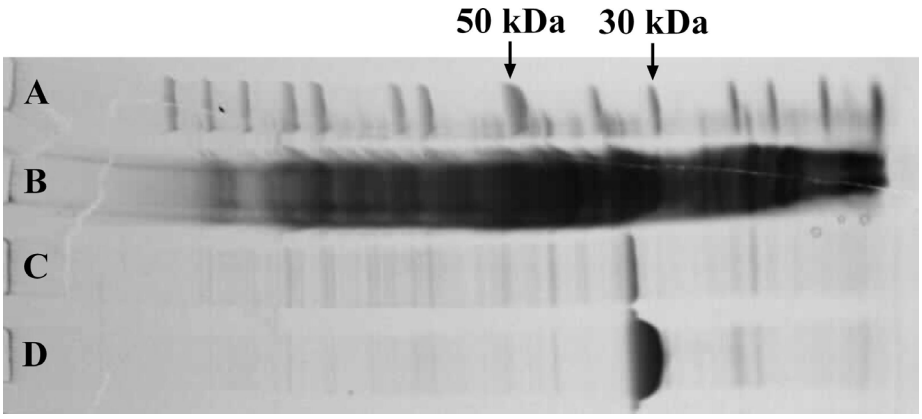


Figure 2.9 – SDS-PAGE of a typical purification of wild-type DHDPS. The lanes were loaded with **(A)** molecular weight marker with relevant bands indicated, **(B)** crude extract, **(C)** sample post IEX and **(D)** sample post HIC.

2.3.2 BIOPHYSICAL CHARACTERISATION

2.3.2.1 KINETICS

The kinetics of DHDPS have been well documented [3, 6, 10, 69]. In this study the kinetic parameters were measured using published procedures [1] and kinetic parameters were obtained by fitting the kinetic data to a pseudo-single substrate model [1, 70]. Values obtained for wild-type DHDPS during the course of this study are consistent with the published values (table 2.1). Kinetic parameters will be used as proxy indicators of an intact active site and correct structure in the variants described in the following chapters.

k_{cat} app (with respect to Pyruvate (s^{-1})) ± SEM	$K_{m\ Pyr}$ app (mM) ± SEM	k_{cat} app (with respect to (*S*)-ASA (s^{-1})) ± SEM	$K_{m\ (S)\text{-}ASA}$ app (mM) ± SEM
118 ± 3	0.32 ± 0.03	123 ± 5	0.14 ± 0.02

Table 2.1 – The kinetic parameters for wild-type DHDPS determined in this study are similar to the published values (k_{cat} ranging from 79 to 188 ± 7 s^{-1}, K_{mPyr} ranging from 0.1 ± 0.01 mM to 0.26 ± 0.03 mM and $K_{m\ (S)\text{-}ASA}$ ranging from 0.11 ± 0.01 mM to 0.13 ± 0.02 mM [3, 6, 10, 69]. The two k_{cat} values arise from the use of pseudo-single substrate models rather than a complete two substrate analysis.

2.3.2.2 CRYSTALLOGRAPHIC ANALYSIS

The crystal structure for wild-type *E. coli* DHDPS was initially obtained by Mirwaldt *et al.* (1995) [71] to a resolution of 2.5 Å and subsequently obtained to a resolution of 1.9 Å by Dobson *et al.*, 2005 [66]. The structure of DHDPS is discussed in chapter 1, section 1.10 and was available as a point of comparison for the mutated DHDPS proteins.

2.3.2.3 ANALYTICAL GEL PERMEATION LIQUID CHROMOTOGRAPHY

Zonal analysis was used to confirm V_e of DHDPS (chapter 8, section 8.6.1). The elution profile resembled a Gaussian distribution, the peak of which was considered the elution volume of the protein [12] (figure 2.10). Molecular mass estimation was achieved through the calibration of the gel filtration column using protein standards of known molecular mass. A calibration plot was constructed using the V_e of these protein standards and from this plot the V_e of wild-type DHDPS was converted to the apparent molecular mass (figure 2.11). Each standard was run in duplicate and was chosen for its similarity to the protein of interest with regard to its Stokes' radius (on which the V_e directly depends) and additionally on its partial specific volume and frictional ratios [8, 12, 72, 73].

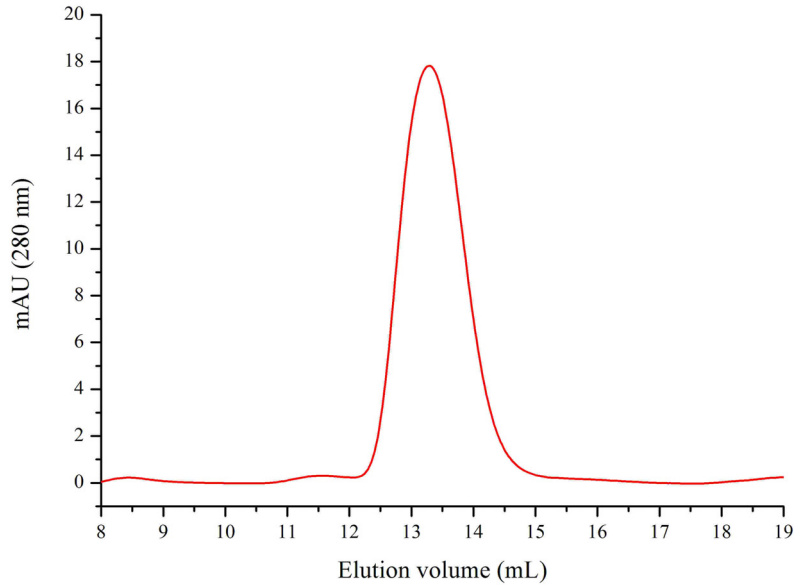


Figure 2.10 – Size exclusion chromatography of wild-type DHDPS on Sephadex G-200 (10/300 GL prepacked) superfine column. The elution volume determined by zonal analysis for wild-type DHDPS (13.3 mL) confirms that it exists as a single oligomeric species in 20 mM Tris (pH 7.0), 100 mM NaCl. The chromatogram was produced by measuring absorbance at 280 nm as a function of elution volume (V_e).

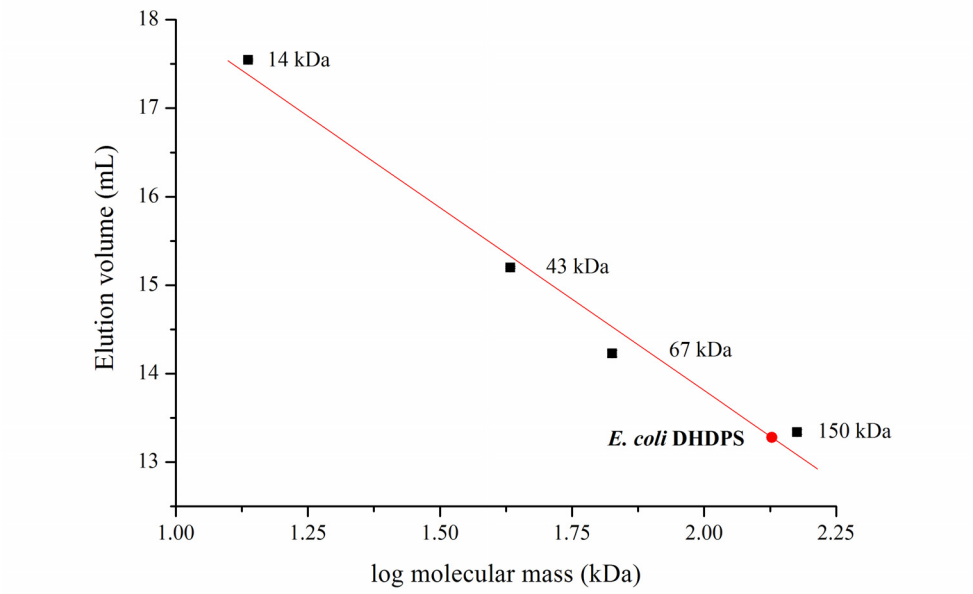


Figure 2.11 – Analytical gel permeation liquid chromatography. The elution volumes of yeast alcohol dehydrogenase (150 kDa), bovine serum albumin (67 kDa), ovalbumin (43 kDa) and ribonuclease A (14 kDa) (in 20 mM Tris (pH 7.0), 100 mM NaCl) plotted against the natural log of their molecular mass. The line of best fit was used to calculate and confirm the molecular mass of wild-type DHDPS.

The predicted molecular mass of the DHDPS tetramer based on its sequence [66, 74] is 125.08 kDa ("Protein Calculator" - http://www.scripps.edu/~cdputnam/protcalc. html). From the calibration plot, the molecular mass of wild-type DHDPS was calculated to be 134 kDa indicating that it exists as a tetramer in solution. The disparity between the predicted molecular mass and that observed is within the 10 % error expected when calculating molecular mass from analytical gel permeation liquid chromatography data [72, 73] and also within 10 % error of the range of molecular mass previously established using gel filtration (112 – 126 kDa) [8, 74, 75].

2.3.2.4 ANALYTICAL ULTRA-CENTRIFUGATION

Sedimentation velocity AUC analysis of DHDPS was performed in order to confirm the thermodynamic and hydrodynamic properties of the protein (chapter 8, section 8.6.2). The assay was performed in 20 mM Tris.HCl (pH 7.0 at 20ºC), 100 mM NaCl, a condition chosen to complement the results of the analytical gel permeation liquid chromatography. The sedimenting boundary did not undergo any significant spreading with respect to time thus suggesting the sample was largely homogeneous (figure 2.12(B)). To confirm this, the data were fitted to a continuous size-distribution model [76, 77]. The $c(s)$ data show a single primary peak with an ordinate maximum which corresponded to a sedimentation coefficient of 6.6 S (figure 2.12(C)). This value is congruent with existing literature for tetrameric DHDPS [8, 10, 78]. The quality of the nonlinear least squares best-fit is affirmed by the low rmsd number of 0.0045 and the random residual distribution (figure 2.12(A)). The small secondary peak with a sedimentation coefficient of 4.2 S, is consistent with previous literature. Perugini *et al.* hypothesised that the smaller species was present as a result of an equilibrium mixture of dimers and tetramers [78]. As SDS-PAGE exhibited no significant proteinaceous contaminants (see figure 2.9), it was concluded that a low level of dissociation of the dimers was occurring, resulting in the presence of the second peak.

The results of the AUC confirm the results obtained by analytical gel permeation chromatography; in solution, DHDPS exists as a tetramer. The corroboration of the AUC and analytical gel permeation chromatography data coupled with the time consuming nature of AUC led to the decision to carry out routine analysis of quaternary structure analysis using analytical gel permeation chromatography and reserve AUC analysis for selected mutants.

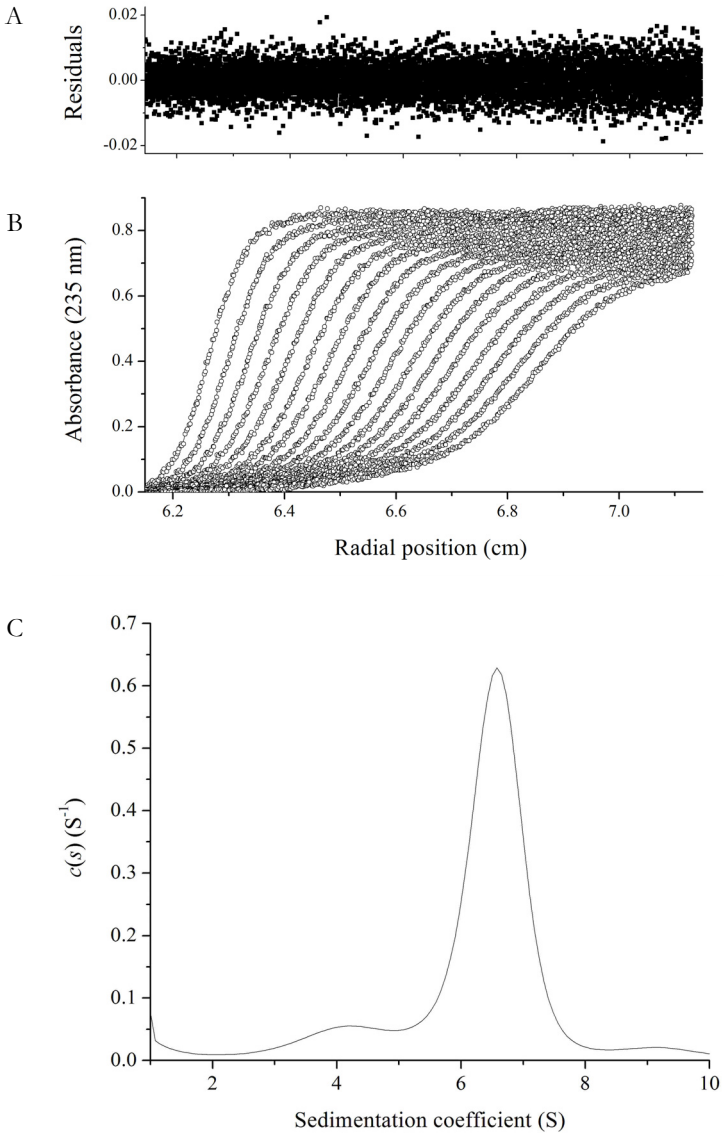


Figure 2.12 – The sedimentation velocity of wild-type DHDPS carried out at 0.1 mg/mL. **(A)** Residuals for the nonlinear least squares best fit are plotted as a function of the radial position from the axis of rotation (cm). **(B)** Absorbance at 235 nm plotted as a function of the radial position from the axis of rotation (cm). Both the raw data (symbols) and fitted data (solid line) are plotted at time intervals of 6 minutes. Raw data are overlaid with the nonlinear least squares best fit to a continuous size-distribution model. **(C)** The continuous sedimentation coefficient ($c(s)$) distribution plotted as a function of the sedimentation coefficient (S). The fit obtained was using a resolution of 100 species between s_{min} of 1 S and s_{max} of 10 S with P = 0.95, $\bar{\nu}$ = 0.7405, ρ = 1.005 g/mL, η = 1.0214 cp and f/f_0 = 1.56.

2.3.2.5 DIFFERENTIAL SCANNING FLUORIMETRY

The use of microplates enabled the screening at pH 2, pH 4.5, pH 7, pH 9.5 and pH 11 with 0 mM, 40 mM, 100 mM and 500 mM NaCl (see section 2.2). Statistical analysis was carried out using a generalised linear model with a Poisson error and a log link function due to the non-normal underlying error distribution. Thermal denaturation temperature was the response variable, and protein species, pH and salt concentrations were the categorical factors. There was no evidence of over-dispersion, and significance tests were made using a Chi-square statistic to a P value of < 0.01. The analysis of deviance table for all variants described in this thesis is presented in appendix 1, table A1.2.

The effect of salt and pH on the thermal denaturation of the wild-type DHDPS in 100 mM phosphate buffer is shown in figure 2.13 and summarised in figure 2.14. The results of this assay suggest salt concentration does not have a large effect on the thermal denaturation temperature of the wild-type DHDPS; however, the effect of pH is significant. At pH 7 and pH 9.5 the protein is the most thermally stable whilst at pH 4.5 and pH 11 the thermal stability is decreased. The data gathered for pH 2 exhibit high initial fluorescence suggesting that the protein is denatured prior to the addition of the dye [17]. The pI of DHDPS in its folded state is 6.57 ("Protein Calculator" - http://www.scripps.edu/ ~cdputnam/protcalc. html), thus pH 7 and pH 9.5 buffers would result in the deprotonation of the protein, increasing solubility and therefore thermal stability [51, 55]. This hypothesis is supported by the fact that the activity of DHDPS is optimal at pH 8.4 indicating a level of deprotonation is required for optimal stability and activity [3, 79]. The reduction in thermal stability at pH 4.5 and pH 11 may be due to the highly protonated and deprotonated state of the protein under these conditions [51].

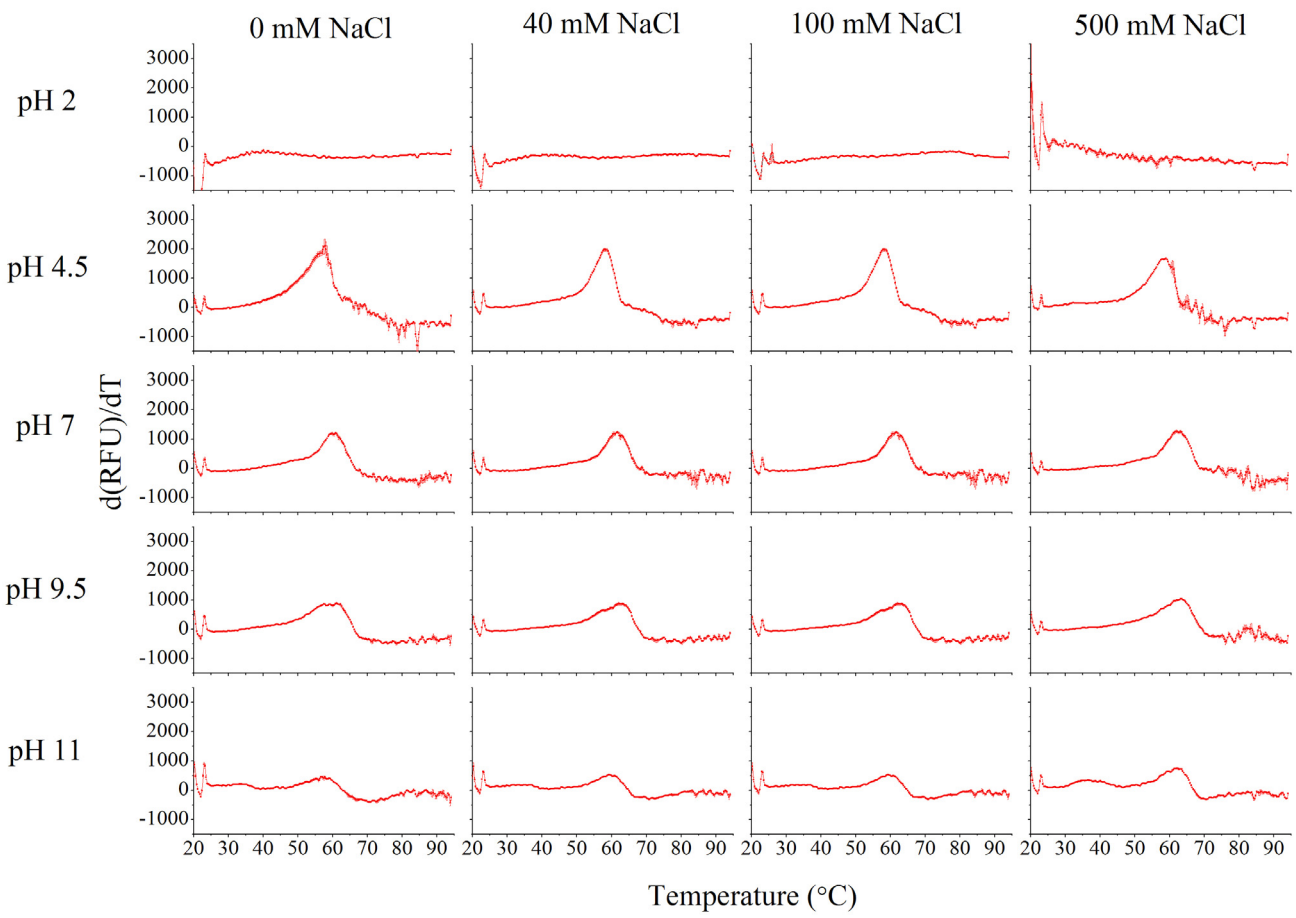


Figure 2.13 – The thermal denaturation profiles of wild-type DHDPS in 100 mM phosphate buffer (pH 7, 40 mM NaCl) as monitored by SYPRO orange fluorescence. The data plotted are the derivatives of the increase in fluorescence monitored by the BioRad IQ5 (presented in appendix A1, figure A1.1). The peaks indicate the temperature at which the protein is unfolding the most rapidly and is recorded as the mean thermal denaturation temperature. The data plotted are the mean values of three replicates ± SEM.

Variable fluorescence signals were observed between buffers at different pH (see appendix A1, figure A1.1). Due to the nature of the technique, such artifacts are a possibility and can be a result of protein aggregation, the presence of impurities and the non-specific interaction of the dye. Most of these would result in an unchanged rate of thermal denaturation but cause variability in fluorescence signal [18]. Thus, despite of the variability observed in the fluorescence signals, the thermal denaturation data for all conditions can be compared.

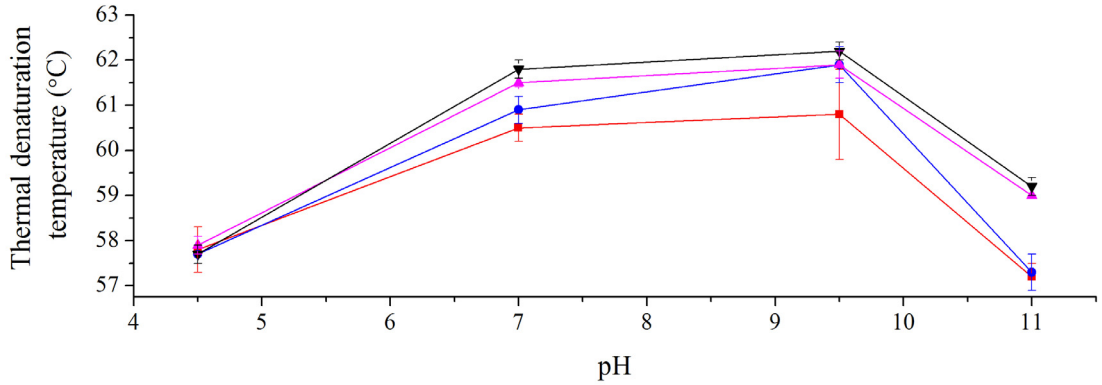


Figure 2.14 – The thermal denaturation temperatures (T_m) of wild-type DHDPS as monitored by SYPRO orange fluorescence in 100 mM phosphate buffer at 0 mM NaCl (red), 40 mM NaCl (blue), 100 mM NaCl (pink) and 500 mM NaCl (black) plotted as a function of pH. The mean thermal denaturation temperatures plotted were calculated from derivations of the fluorescence data. The error bars represent the SEM.

The results of the DSF assays indicate that NaCl has only a small role in the determination of the thermal denaturation temperature of DHDPS. The effect of pH on thermal denaturation temperature is significant and most likely reflects the state of protonation or deprotonation of the proteins at different pH.

DSF used in the manner detailed above provides information regarding the thermal stability of proteins under a range of conditions but is limited by its mechanism of detection of hydrophobic patches. In order to corroborate the results of the thermal denaturation assay, circular dichroism spectroscopy was performed.

2.3.2.6 CIRCULAR DICHROISM SPECTROSCOPY

The CD spectrum of the wild-type enzyme in 20 mM phosphate buffer (pH 7), 40 mM NaCl at 20 ºC was determined in order to provide baseline data with regard to the secondary structure composition of DHDPS. This buffer was chosen because all the assays performed under this condition assessing stability (sections 2.3.2.5, 2.3.2.7 and 2.3.2.8) yielded quantifiable results enabling comparative analysis. Figure 2.15 shows the spectrum obtained from a wavelength scan. It

demonstrates that DHDPS is folded in solution, exhibiting double minima at 208 nm and 222 nm, indicative of predominantly α-helical structures, including $(\alpha/\beta)_8$ barrels [80-82] and consistent with the spectra previously obtained for wild-type DHDPS [4, 9, 10].

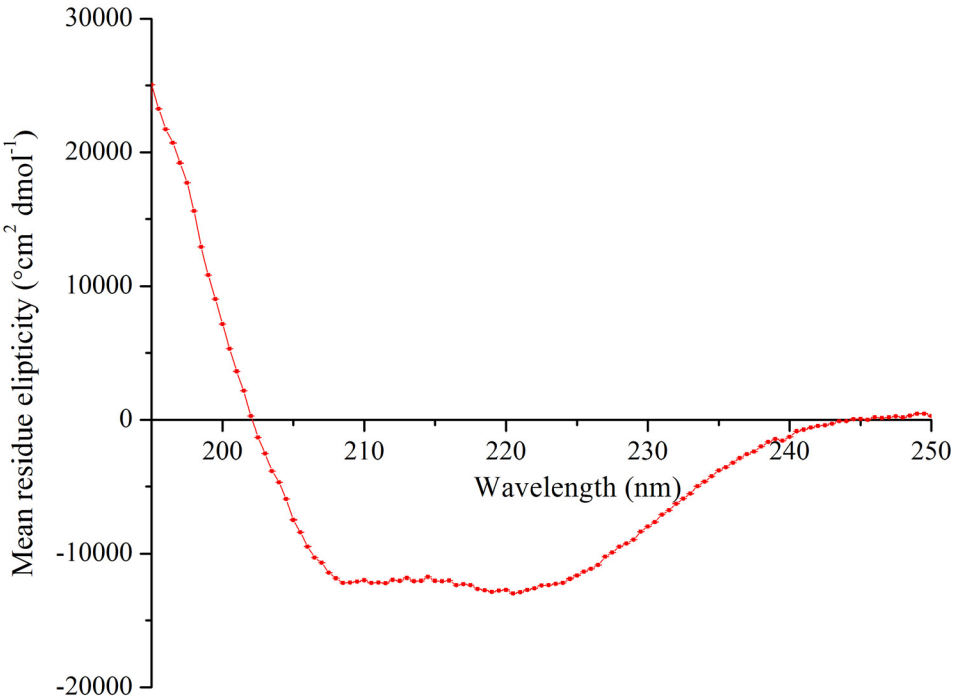


Figure 2.15 – Circular dichroism spectroscopy of DHDPS (at 0.2 mg/mL in 20 mM phosphate buffer (pH 7), 40 mM NaCl). Mean residue ellipticity is plotted against wavelength.

The secondary structure of proteins can be quantified from wavelength scan data using a range of algorithms [23, 26, 83, 84]. These are not as accurate as the results obtained from crystallographic analysis [23] (table 2.2), but do provide a point of comparison for solution structures that is useful if crystal structures are not available.

	Helix	**Sheet**	**Other**
Crystallographic data	56 %	12 %	32 %
CD analysis	41 % ± 1 %	14 % ± 1 %	45 % ±1 %

Table 2.2 – The secondary structure composition of DHDPS according to the crystallographic data obtained by Dobson *et al.*, 2004 [66] (PDB accession number 1yxc), and analysis of CD data obtained during this research using CDPro [26]. The data obtained from CDPro are the means of the values obtained from CDSSTR, CONTILL and SELCON3 using reference set 4.

The changes in secondary structure of DHDPS were assessed with increasing temperature. This was achieved by monitoring the change in intensity at 220 nm, indicative of a change in the proportion of α-helix present in the protein. The dissolution of the secondary structure or thermal denaturation can be used to calculate the "melting temperature" of the protein. During the process of thermal denaturation of wild-type DHDPS, wavelength scans were performed every 10ºC between 20ºC and 80ºC in order to assess the unfolding "landscape" [24]. This was carried out three times to estimate reproducibility (figure 2.16). An unsuccessful attempt was made to refold DHDPS through a reversal of the thermal denaturation process. These data are presented in appendix 1, section A1.2.

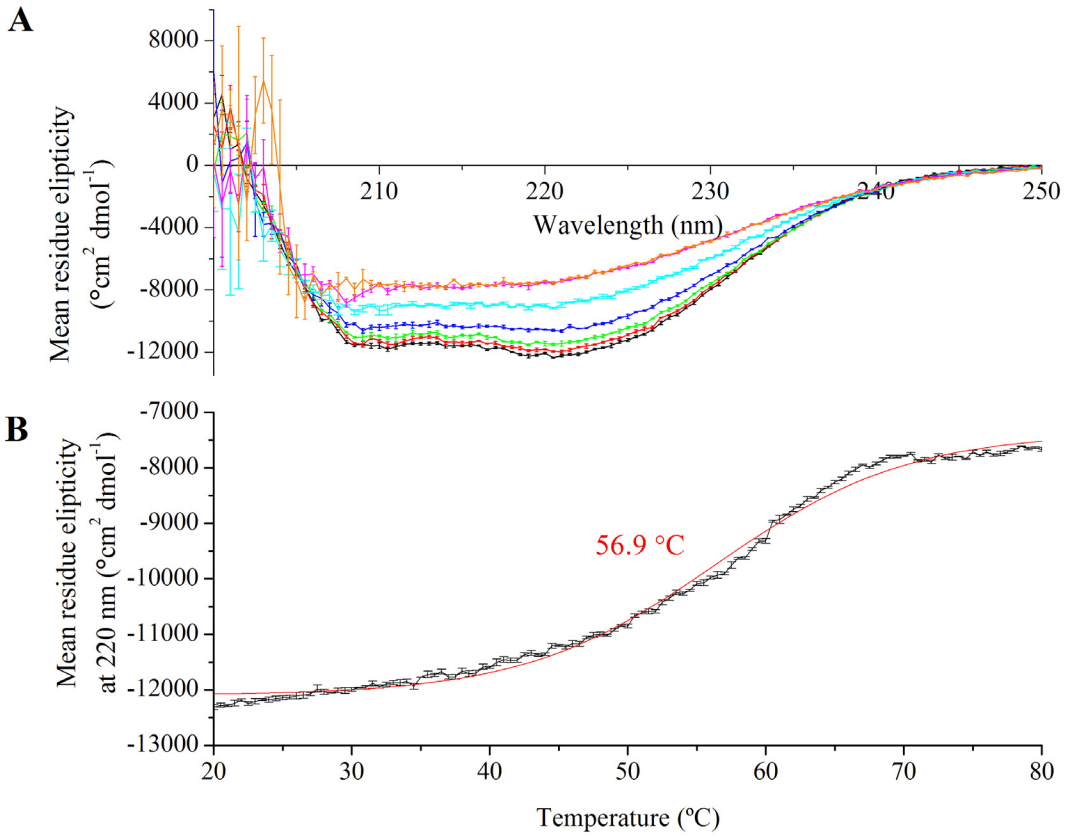


Figure 2.16 – Circular dichroism spectroscopy of wild-type DHDPS (at 0.05 mg/mL in 20 mM phosphate buffer (pH 7.0), 40 mM NaCl). **A)** wavelength scans plotting mean residue ellipticity [θ] vs. wavelength at 20ºC (black), 30ºC (red), 40ºC (green), 50ºC (dark blue), 60ºC (light blue), 70ºC (pink) and 80ºC (orange). **B)** Mean residue ellipticity at 220 nm vs. temperature indicating melting of the protein. The red line corresponds to a sigmoid (Boltzmann) fit and the thermal denaturation temperature shown (in red) was calculated by identifying the midpoint of the transition as previously described [17]. The data plotted are the mean values of three replicates and the error bars represent the SEM.

The change in secondary structural composition corresponding to increasing temperature can be calculated using algorithms [23, 26]. This analysis was not performed because none of the algorithms are accurate with regard to β-sheet rich proteins thus would not provide data as to whether DHDPS was enriched for β-sheet as a function of temperature [23]. In addition to this, the decline in signal intensity upon increasing temperature indicates that the concentration of DHDPS in solution is decreasing. This is attributable to aggregation and would compromise any structural analysis performed.

Error values of the thermal denaturation data of wild-type DHDPS were calculated from three replicate experiments. The error of the melt curves was calculated by fitting sigmoid (Boltzmann) distributions to each of the three replicates (figure 2.17). The midpoints of each fit were averaged and the SEM calculated.

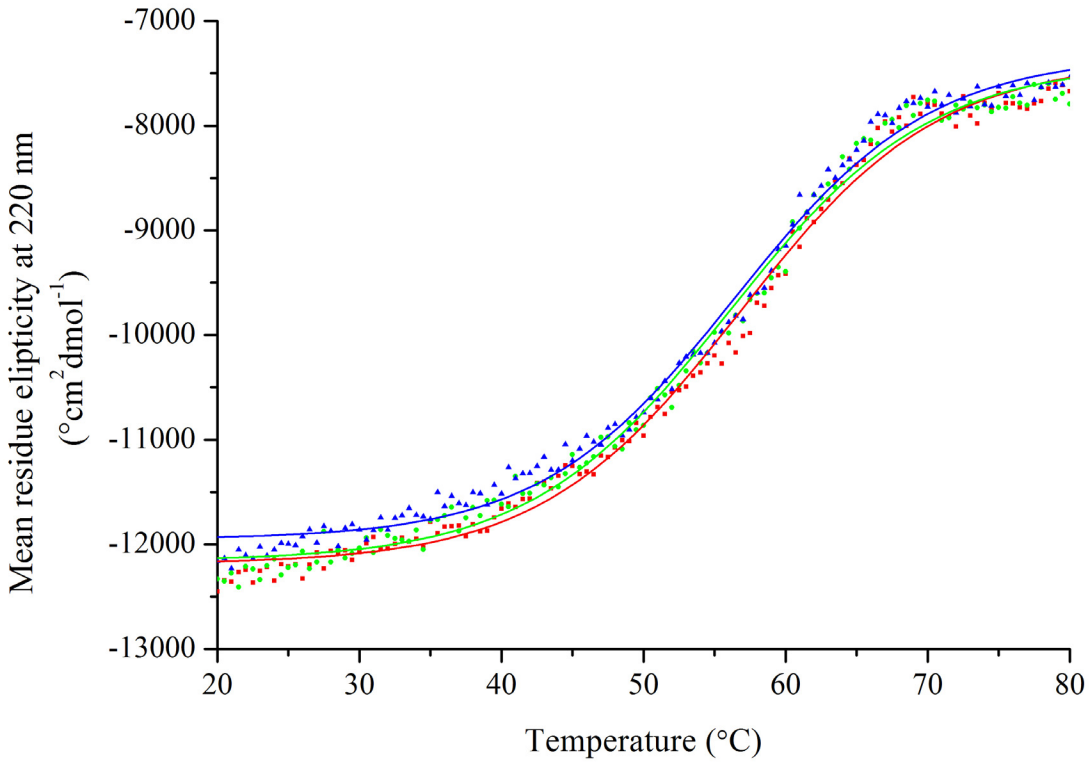


Figure 2.17 – Thermal denaturation data of wild-type DHDPS monitored by CD spectroscopy. Mean residue elipticity at 220 nm vs. temperature is plotted. The red, green and blue data points represent the three replicates carried out. The lines in the respective colours correspond to a sigmoid (Boltzmann) fit of each replicate and exhibited errors of < 3%. The temperatures used to calculate the mean denaturation temperature were determined by identifying the midpoint of the transition as previously described [17].

The mean temperature calculated was 56.9 ± 0.15 ºC which is slightly lower than the thermal denaturation temperature previously reported by Pearce *et al.*, 2006 and 2008 [9, 85] although the conditions under which this assay was performed were different. Those published by Pearce *et al.* were performed at pH 8 and those carried out during the course of this thesis were performed at pH 7. As pH has been shown to affect the thermal denaturation of DHDPS (section 2.3.2.5), this variability in thermal denaturation temperatures is not surprising. Due to the low error, the relatively time consuming nature of the experiments and the large number of proteins, it was concluded that only one thermal denaturation profile per protein was necessary for the variants described in the following chapters.

The effect of pH on wild-type DHDPS was investigated. The wavelength scans (performed in 20 mM phosphate buffer supplemented with 40 mM NaCl) indicated that in the presence of salt, the secondary structure of the protein is similar at pH 2, pH 4.5, pH 7, pH 9.5 and pH 11 (appendix 1, section A1.2). The DSF analysis (section 2.3.2.5) demonstrated high initial fluorescence at pH 2, indicative of the presence of many exposed hydrophobic groups. These two facts led to the hypothesis that at high and low pH (correlating to increased levels of deprotonation or protonation), secondary structure is maintained, however tertiary structure is disrupted. This is supported by the fact that DSF (section 2.3.2.5) and CD (appendix A1, section A1.2) indicate that DHDPS thermally denatures at a significantly lower temperature at pH 4.5 and pH 11 than at pH 7 or pH 9.5; the extensive protonation and deprotonation at pH 4.5 and pH 11 reduces the stability of the protein by weakening the tertiary structure.

The thermal denaturation temperatures of wild-type DHDPS as determined by DSF and CD (in 100 mM phosphate buffer (pH 7), 40 mM NaCl) were 60.9 ± 0.3 and 57.0 ± 0.1 ºC respectively. There is variability observed between the CD and DSF techniques, consistent with the literature on other proteins [18, 86, 87]. The results of both techniques were qualitatively similar indicating a thermal denaturation temperature of approximately 60 ºC in pH 7, 40 mM NaCl. The variability observed between techniques was attributed to the underlying natures of each of the methods. Whilst DSF measures the binding and subsequent fluorescence of the SYPRO orange dye upon exposure of hydrophobic patches, CD spectroscopy measures the dissociation of the secondary structure. It was hypothesised that CD yielded lower thermal denaturation temperatures than DSF because during the unfolding of DHDPS the secondary structure is lost before hydrophobic residues are exposed. Since the DSF technique correlated well with CD and allowed high throughput screening, it was routinely

used to screen all mutants across the full range of conditions in later experiments, with limited CD spectra to corroborate these results.

2.3.2.7 AMORPHOUS AGGREGATION

Amorphous aggregate formation (see figure 2.7) was quantified through monitoring the increase in light scattering at 340 nm [28]. The rate and extent of aggregation of wild-type DHDPS was screened across the full range of conditions (see section 2.2). The assays were performed at 60 ºC, the midpoint of the thermal denaturation of wild-type DHDPS according to DSF and CD spectroscopy. In theory, the midpoint of the thermal denaturation curve represents the point at which half the protein is folded and the other half has aggregated. The presence of the partially folded protein population was assumed to facilitate the formation of aggregates, and β-sheet-specific structures [11, 51, 88, 89]. Again, the use of micro-plates enabled the screening of all the conditions simultaneously (presented in their entirety in appendix 1, section A1.3). A typical selection is presented here.

In order to measure the speed of amorphous aggregation, sigmoid functions were fitted to the aggregation data. The aggregation curves were characterised by a lag period followed by a rapid increase in light scattering. Once the light scattering has reached its maximum, the linear phase begins. This linear phase is characterised by little or no increase in light scattering with respect to time. The sigmoid function was only fitted up until this point due to a progressive decrease in light scattering attributed to the sedimentation of the protein aggregate [16] and exhibited a confidence interval of greater than 98 % for all data analysed. Only the data obtained for pH 7 and pH 9.5 were conducive to this analysis. The other pH conditions (pH 2, pH 4.5 and pH 11) resulted in extremely rapid aggregation or demonstrated no light scattering. The rapid precipitation was attributed to the protein being insoluble in the buffer in which it was incubated. The absence of any light scattering could be attributable to the protein remaining in solution at pH 2 and pH 11, a hypothesis supported by visual assessment of the assay which revealed no visible particulate or aggregated material [90, 91]. The results of the CD analysis of DHDPS at pH 2 and pH 11 revealed that they can exist in folded or unfolded states dependent on the presence or absence of salt (see appendix 1, figure A1.6). Because the results of the light scattering assays were similar over the range of salt concentrations tested, it was concluded that whether the protein is folded does not have any effect on the aggregation at either pH 2 or pH 11. This is most likely due to the highly charged states of the

protein resulting in a reduced propensity for inter- and intra-molecular aggregation. The results regarding the effect of salt and pH on amorphous aggregation are summarised in figure 2.18 and appendix 1 (section A1.3).

In order to quantify the aggregation of protein, two parameters were measured. The first was a measure of the time it took for half of the aggregate to form, achieved through the calculation of the midpoint ($t_{½}$) of the sigmoid curve fitted to aggregation data [92]. This technique is a method described by Sabate *et al.*, 2003 [92] and allowed a quantification of aggregation kinetics, enabling comparative analysis. This parameter was termed $Agg_{½}$ (expressed in minutes) and is referred to as the aggregation half life throughout this thesis. The second parameter measured was the maximum light scattering exhibited during the initial 90 minutes of the assay. Statistical analysis was carried out using a multifactor ANOVA, with pH, and salt as categorical factors. Significance is presented to a P value of < 0.01. The analysis of deviance table for $Agg_{½}$ is presented in appendix 1, table A1.4. Only the results at pH 7 and pH 9.5 were used in this statistical analysis because $Agg_{½}$ was not able to be calculated at other pH values due the inability to fit a sigmoid curve to the data. The analysis of deviance table for maximum light scattering is presented in appendix 1, table A1.5.

The effects of salt and pH on the aggregation of wild-type DHDPS in 100 mM phosphate buffer is shown in figure 2.18 and demonstrates that both have a significant effect. The data for maximum light scattering is summarised in figure 2.19. At pH 7, low salt concentrations (0 mM, 20 mM, 40 mM and 100 mM NaCl) did not affect the $Agg_{½}$; however, at 500 mM NaCl, there was a significant decrease in aggregation half life. This probably reflects a decreased solubility brought about by the reduction of the solvating power of NaCl at relatively high concentrations [55]. At pH 9.5, the effect of salt was significant at all concentrations. The maximum light scattering exhibited by DHDPS was significantly affected by the presence of salt with variable effects dependent on pH. Increasing salt concentrations generally caused an increase in the maximum light scattering, except for at pH 7 which exhibited the opposite trend.

The aggregation of the wild-type DHDPS is significantly affected by pH. This result is entirely expected as it is an established fact that pH has a major influence on protein stability and therefore aggregation [46, 55, 90, 91, 93-95]. The aggregation of wild-type DHDPS was the most rapid at pH 4.5 and decreased with increasing pH. The maximum light scattering increased between pH 4.5 and pH 9.5 whilst at pH 2 and pH 11 no light scattering was observed (except for with 500 mM NaCl).

This may have been due to the improved solubility of the proteins under these conditions, a result of the protonated and deprotonated state of the protein respectively. The hypothesis that the protein is prevented from forming aggregate as a result of high net charge on the surface and not the unfolded nature of the protein is supported by the fact that CD spectroscopy indicated that some secondary structure is retained over a wide range of pH values (section 2.3.2.6 and appendix A1, section A1.2).

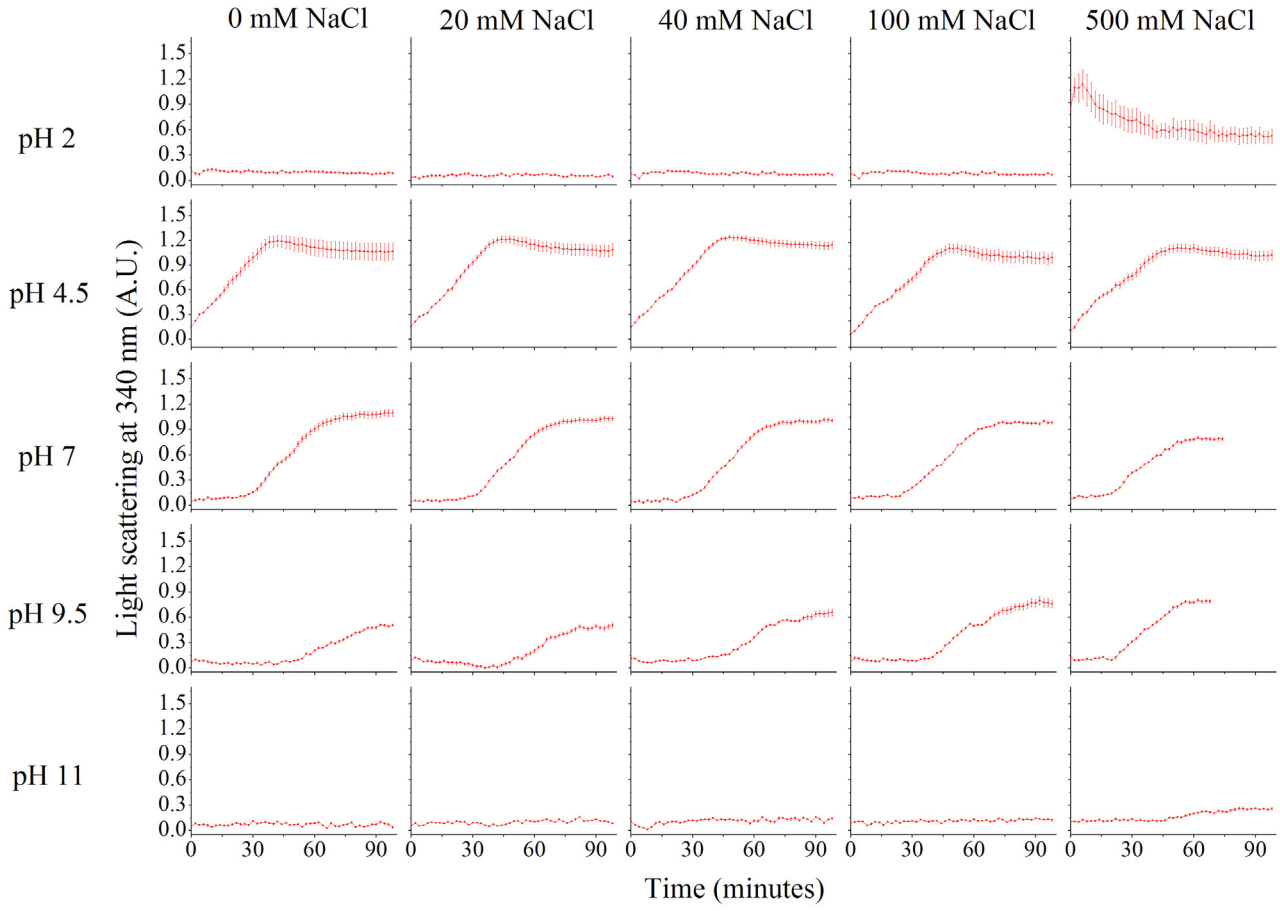


Figure 2.18 — The aggregation propensity of wild-type DHDPS in 100 mM phosphate buffer. Where possible, a sigmoid (Boltzmann) function was used for the line of best fit (> 98% confidence interval) in order to calculate the aggregation half life ($Agg_{½}$). The $Agg_{½}$ was determined by identifying the midpoint of the transition as previously described [92]. Only the data up until the linear phase was included. Once the peak turbidity was reached the light scattering was inconsistent due to the sedimentation of the protein aggregate. The maximum light scattering was the highest reading over the first 90 minutes of the assay. The data plotted are the mean values of six replicates and the error bars represent the SEM.

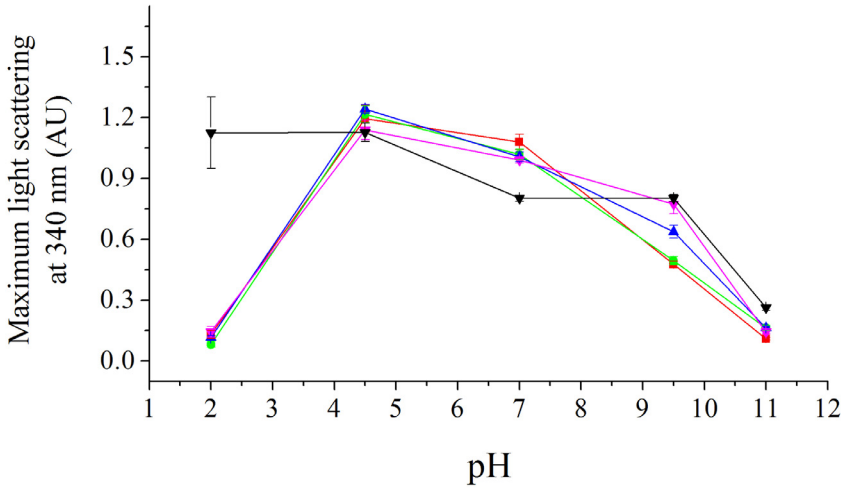


Figure 2.19 –The aggregation of wild-type DHDPS in 0 mM NaCl (red), 20 mM NaCl (green), 40 mM NaCl (blue), 100 mM (pink) and 500 mM NaCl (black). The data plotted are the maximum absorbance units reached over 90 minutes and are plotted as a function of pH. The error bars represent the SEM.

2.3.2.8 β-SHEET-SPECIFIC AGGREGATION

In addition to amorphous aggregate formed by DHDPS (section 2.3.2.7), β-sheet-specific aggregates were investigated. β-Sheet specific aggregates are of interest as they can be indicative of amyloid formation.

The β-sheet-specific aggregation propensities of wild-type DHDPS were screened in 100 mM phosphate buffer at pH 2, pH 4.5, pH 7, pH 9.5 and pH 11 each containing 0 mM, 40 mM, and 500 mM NaCl and in the presence of 1.6 mM ThT. The assays were carried out for 72 hours; however, no further increase in fluorescence was seen after approximately 7 hours. The β-sheet specific aggregation half life ($\beta\text{-}agg_{½}$) was calculated from the sigmoid (Boltzmann) curve fitted to the fluorescence data, by identifying the $t_{½}$ in an analogous manner to that described for the amorphous aggregate data (section 2.3.2.7) [92] and exhibited confidence intervals of greater that 98 % for all data analysed. The maximum fluorescence reading during the initial 300 minutes of the assay was also measured. These data were used in the comparison of the β-sheet specific aggregation propensity of the variants described in the following chapters.

The fluorescence resulting from excitation of the samples at 450 nm was monitored at 482 nm (emission). An increase in fluorescence indicates the formation of β-sheet structures, a result which may be indicative of amyloid fibril formation [32, 33]; however, non-specific binding of ThT has been reported as a result of protein composition and structure [37]. A sample of insulin was incubated at pH 1.7 in parallel with wild-type DHDPS as a positive control (see chapter 8, section 8.6.6). This control sample routinely exhibited a fluorescence in excess of 65 000 RFU indicative of the strongly amyloidogenic nature of insulin [96, 97].

The ThT induced fluorescence of the wild-type DHDPS in 100 mM phosphate buffer is shown in figure 2.20 and comprehensive results presented in appendix 1, section A1.4. The data obtained exhibited a distribution characteristic of ThT fluorescence often attributed to the kinetics of amyloid fibril formation [32-36]. Quantification of the β-sheet specific aggregation half life was carried out on pH 2, 500 mM NaCl, pH 4.5 and pH 7 only. The results obtained for pH 2, 0 mM NaCl and 40 mM NaCl, pH 9.5 and pH 11 were not conducive to the Boltzmann analysis due to low or inconsistent fluorescence (figure 2.20). The maximum fluorescence is summarised in figure 2.21. The results of the amorphous aggregation assays (section 2.3.2.7) and the CD analysis (section 2.3.2.6) would suggest that DHDPS is soluble at pH 2 and pH 11, explaining the lack of fluorescence. The low fluorescence at pH 9.5 may have been due to the formation of aggregates containing few β-sheet structures.

The effects of pH on the ThT fluorescence in the presence of wild-type DHDPS cannot be compared as it has been shown that ThT fluorescence is pH dependent in that, although the capacity for binding is the same [37, 98], the affinity of ThT for β-sheet structures is variable [33], affecting the $\beta\text{-}agg_{½}$. The effect of salt on ThT fluorescence was tested using insulin, a protein with a well documented propensity for forming amyloid fibrils. The results of these assays demonstrated that salt significantly altered both the kinetics of fibril formation and the binding of ThT (see appendix 1, figure A1.8), consistent with the results published by Nielsen *et al.* who found that ionic strength affects the kinetics of fibril formation [98]. For this reason, the effect of salt on the fluorescence of ThT in the presence of DHDPS cannot be compared; however, the data obtained for wild-type DHDPS are used in the comparative analysis of other DHDPS variants.

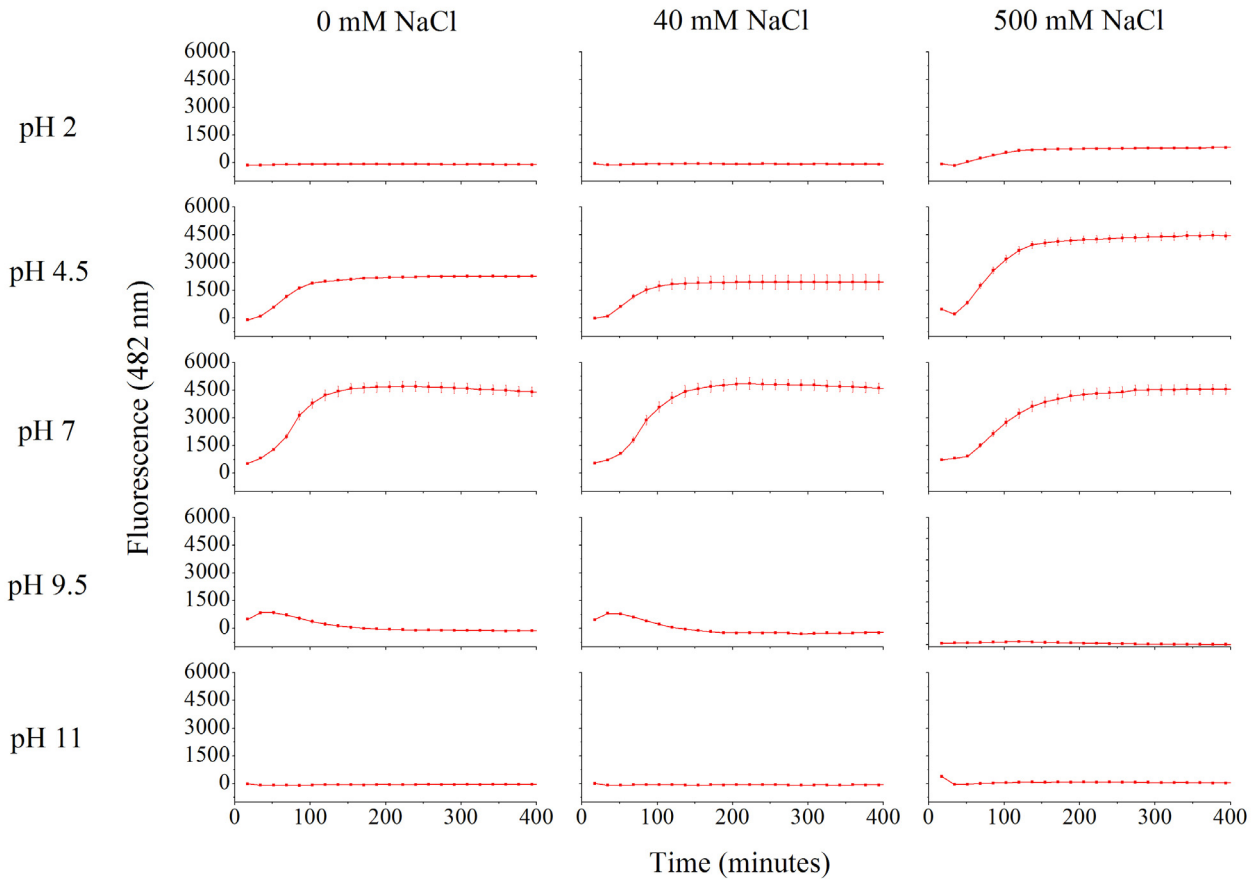


Figure 2.20 – The ThT induced fluorescence of DHDPS in 100 mM phosphate buffer (pH 7) at 40 mM NaCl. Note that only the data up until the linear phase was included [16]. Once the maximum fluorescence was reached the fluorescence readings were inconsistent due to the precipitation of the protein aggregate [16]. The data plotted are the mean values of six replicates and the error bars represent the SEM. Where possible, a sigmoid (Boltzmann) function was used for the line of best fit (> 98% confidence interval) in order to calculate the β-sheet specific aggregation half life (β-$agg_{½}$). The β-$agg_{½}$ was determined by identifying the midpoint of the transition as previously described [92]. The maximum fluorescence was the highest reading over the first 300 minutes of the assay.

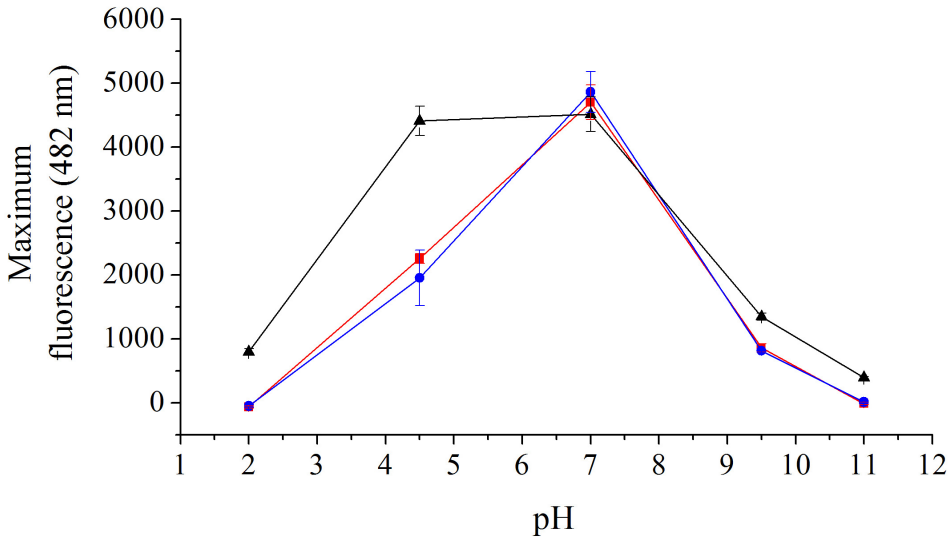


Figure 2.21 – The aggregation of wild-type DHDPS in 0 mM NaCl (red), 40 mM NaCl (blue), and 500 mM NaCl (black). The data plotted are the maximum fluorescence reached over 300 minutes and are plotted as a function of pH. The error bars represent the SEM.

2.3.2.9 CONFIRMATION OF AMYLOID FIBRILS

Transmission electron microscopy (TEM) and X-ray fibre diffraction were utilised to corroborate the ThT results with regard to wild-type DHDPS. The confirmation of the presence of amyloid fibrils in samples prepared under the same conditions as those that exhibited ThT induced fluorescence was attempted using TEM and X-ray fibre diffraction (chapter 8, section 8.8). Despite extensive trials, no fibrillar structures were observed under TEM (see figure 2.22) on any of the grids examined. The inability to confirm the presence of amyloid in those samples that exhibited ThT fluorescence suggests that β-sheet structures may be forming, but they are not true amyloid fibrils. This is consistent with the fact that no $(\alpha/\beta)_8$ barrel has been shown to form an amyloid fibril and that ThT fluorescence in the presence of DHDPS was much lower than that of insulin. Attempts to obtain X-ray fibre diffraction stalks were unsuccessful because the surface tension of protein suspension (chapter 8, section 8.8.2) was not strong enough to maintain the drop between the capillaries during the drying process.

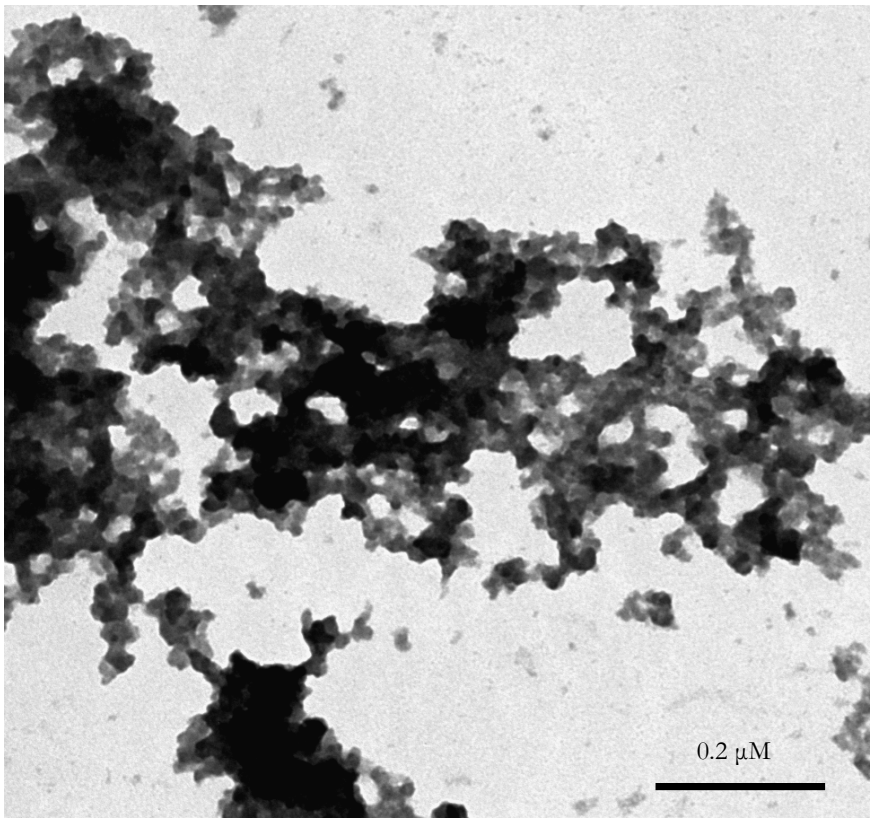


Figure 2.22 – TEM image of non-specific aggregate formed by wild-type DHDPS at 0.5 mg/mL in 40 mM NaCl at pH 7. This image was typical of all 10 images obtained.

2.4 SUMMARY AND CONCLUSIONS

A range of biochemical techniques described in this chapter have been used to ascertain and/or confirm the biophysical properties of *E. coli* DHDPS and provide reference data for subsequent chapters in which the same techniques were utilised to investigate the properties of DHDPS variants.

Kinetic analyses have been used as a proxy indicator of native fold, because activity of DHDPS is dependent on both tertiary and quaternary structure. Crystallography can be used to corroborate the results obtained from kinetic analyses by providing accurate information regarding all levels of structure and organisation, but only in the crystal form. The quaternary structure of proteins has been assessed in solution using analytical gel permeation liquid chromatography and/or analytical ultracentrifugation. Analytical gel permeation liquid chromatography is more convenient than

crystallography and AUC, thus has been used extensively in the following chapters. The thermal stability and folding of proteins has been investigated using DSF and CD spectroscopy. The propensity of the protein to form amorphous aggregate was another indicator of stability. The increase in fluorescence of DHDPS upon incubation in the presence of thioflavin T was an indication of the presence of β-sheet structures; however, further TEM and X-ray fibre showed no evidence of amyloid fibrils

The results of each of the biophysical assays described in this chapter are interrelated. For example, the temperature at which a protein denatures has a direct bearing on the temperature at which amorphous and β-sheet-specific aggregates are formed. In the case of wild-type DHDPS the thermal denaturation temperature (~60 ºC according to DSF and CD (sections 2.3.2.5 and 2.3.2.6)) was used to promote the formation of amorphous and β-sheet-specific aggregates. The thermal denaturation temperature exhibited by DSF did not vary greatly with salt concentration. The reports regarding the influence of salt on the formation amorphous and β-sheet-specific aggregates supports this [50, 64]. The variability in thermal denaturation temperature with regard to pH is diagnostic of altered stability. This in turn leads to variable propensities to form amorphous aggregate and, as the occurrence of aggregation is thought to be a critical step in the process of protofilament formation (see figure 1.1), it may also be affecting the kinetics of β-sheet-specific aggregate formation. As there appears to be little correlation between the thermal denaturation temperature according to the results of the DSF and the amorphous and β-sheet-specific aggregation propensity of wild-type DHDPS, it is reasonable to speculate that the variability observed in the $Agg_{½}$ and $\beta\text{-}agg_{½}$ is attributable to the buffer conditions in which the protein was incubated.

In general, the effect of salt on the thermal denaturation temperature, amorphous aggregation and β-sheet-specific aggregation of wild-type DHDPS was minimal apart from at high (500 mM) NaCl. The decreased $Agg_{½}$ at 500 mM NaCl correlates with an increased $\beta\text{-}agg_{½}$. Since the equilibria involved in the formation of amorphous and β-sheet-specific aggregation are in competition, when the formation of one type of aggregation is favoured, the formation of the other is inhibited. Alternatively the more rapid formation of amorphous aggregate leads to the formation of particulate species that are not conducive to the formation of the precursors of β-sheet-specific aggregate [44].

The effect of pH plays a significant role in determining the thermal denaturation temperatures, amorphous aggregation and β-sheet-specific aggregation. With the exception of pH 2 and pH 11, it

appears that an increased thermal denaturation temperature correlates with formation of amorphous aggregate and decreases the maximum light scattering. The influence of pH on ThT fluorescence cannot be compared (as discussed in section 2.3.2.8). Nevertheless, the results of the ThT assays at pH 4.5 and pH 7 are used in the comparison of the DHDPS variants described in the following chapters.

Using the biophysical techniques described in this chapter, an investigation has been carried out into the effect of temperature, pH and salt on the equilibria of the folding pathway of wild-type DHDPS. Using the results obtained, comparative analysis of the variants described in the following chapters will be carried out.

2.5 REFERENCES

1 Cornish-Bowden, A. (1999) Fundamentals of enzyme kinetics. Portland Press Ltd., London

2 Coulter, C. V., Gerrard, J. A., Kraunsoe, J. A. E. and Pratt, A. J. (1999) *Escherichia coli* dihydrodipicolinate synthase and dihydrodipicolinate reductase: Kinetic and inhibition studies of two putative herbicide targets. Pesticide Science **55**, 887 - 895

3 Karsten, W. E. (1997) Dihydrodipicolinate synthase for *Escherichia coli*: pH dependent changes in the kinetic mechanism and kinetic mechanism of allosteric inhibition by *L*-lysine. Biochemistry **36**, 1730 - 1739

4 Dobson, R. C. J., Devenish, S. R. A., Turner, L. A., Clifford, V. R., Pearce, F. G., Jameson, G. B. and Gerrard, J. A. (2005) Role of arginine 138 in the catalysis and regulation of *Escherichia coli* dihydrodipicolinate synthase. Biochemistry **44**, 13007 - 13013

5 Dobson, R. C. J., Gerrard, J. A. and Pearce, F. G. (2004) Dihydrodipicolinate synthase is not inhibited by its substrate, (*S*)-aspartate β-semialdehyde. Biochemical Journal **377**, 757 - 762

6 Dobson, R. C. J., Griffin, M. D., Roberts, S. J. and Gerrard, J. A. (2004) Dihydrodipicolinate synthase (DHDPS) from *Escherichia coli* displays partial mixed inhibition with respect to its first substrate, pyruvate. Biochimie **86**, 311 - 315

7 Paiva, A. M., Vanderwall, D. E., Blanchard, J. S., Kozarich, J. W., Williamson, J. M. and Kelly, T. M. (2001) Inhibitors of dihydrodipicolinate reductase, a key enzyme in the diaminopimelate pathway of *Mycobacterium tuberculosis*. Biochimica et Biophysica Acta **1545**, 67 - 77

8 Griffin, M. D. W. (2005) PhD Thesis. Why is DHDPS a tetramer? University of Canterbury, Christchurch

9 Pearce, F. G., Dobson, R. C., Weber, A., Lane, L. A., McCammon, M. G., Squire, M. A., Perugini, M. A., Jameson, G. B., Robinson, C. V. and Gerrard, J. A. (2008) Mutating the tight-dimer interface of dihydrodipicolinate synthase disrupts the enzyme quaternary structure: towards a monomeric enzyme. Biochemistry **In press**

10 Griffin, M. D., Dobson, R. C. J., Pearce, F. G., Antonio, L., Whitten, A. E., Liew, C. K., Mackay, J. P., Trewhella, J., Jameson, G. B., Perugini, M. A. and Gerrard, J. A. (2008)

Evolution of quaternary structure in a homotetrameric enzyme. Journal of Molecular Biology **380**, 691 - 703

11 Dobson, C. M. (2004) Principles of protein folding, misfolding and aggregation. Seminars in Cell and Developmental Biology **15**, 3 - 16

12 Winzor, D. J. (2003) Analytical exclusion chromatography. Journal of Biochemical and Biophysical Methods **56**, 15 - 52

13 Laue, T. M. and Stafford, W. F. (1999) Modern applications of analytical ultracentrifugation. Annual Review of Biophysics and Biomolecular Structure **28**, 75 - 100

14 Lebowitz, J., Lewis, M. S. and Schuck, P. (2002) Modern analytical ultracentrifugation in protein science: a tutorial review. Protein Science **11**, 2067 - 2079

15 Vistica, J., Dam, J., Balbo, A., Yikilmaz, E., Mariuzza, R. A., Rouault, T. A. and Schuck, P. (2004) Sedimentation equilibrium analysis of protein interactions with global implicit mass conservation constraints and systematic noise decomposition. Analytical Biochemistry **326**, 234 - 256

16 Niesen, F. H., Berglund, H. and Vedadi, M. (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. Nature Protocols **2**, 2212 - 2221

17 Vedadi, M., Niesen, F. H., Allali-Hassani, A., Fedorov, O. Y., Finerty, P. J. J., Wasney, G. A., Yeung, R., Arrowsmith, C., Ball, L. J., Berglund, H., Hui, R., Marsden, B. D., Norlund, P., Sundstrom, M., Weigelt, J. and Edwards, A. M. (2006) Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. Proceedings of the National Academy of Sciences USA **103**, 15835 - 15840

18 Epps, D. E., Sarver, R. W., Rogers, J. M., Herberg, J. T. and Tomich, P. K. (2001) The ligand affinity of proteins measured by isothermal denaturation kinetics. Analytical Biochemistry **292**, 40 - 50

19 Senisterra, G. A., Hong, B. S., Park, H.-W. and Vedadi, M. (2008) Application of high-throughput isothermal denaturation to assess protein stability and screen for ligands. Journal of Biomolecular Screening **13**, 337 - 342

20 Steinberg, T. H., Haugland, R. P. and Singer, V. L. (1996) Applications of SYPRO orange and SYPRO red protein gel stains. Analytical Biochemistry **239**, 238 - 245

21 Steinberg, T. H., Jones, L. J., Haugland, R. P. and Singer, V. L. (1996) SYPRO orange and SYPRO red gel stains: One step fluorescent staining of denaturing gels for detection of nanogram levels of protein. Analytical Biochemistry **239**, 233 - 237

22 Stryer, L. (1965) The interaction of a napthalene dye with apomyoglobulin and apohemoglobin. A fluorescent probe of non-polar binding sites. Journal of Molecular Biology **13**, 482 - 495

23 Greenfield, N. J. (2007) Using circular dichroism spectra to estimate protein secondary structure. Nature Protocols **1**, 2876 - 2890

24 Greenfield, N. J. (2007) Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. Nature Protocols **1**, 2527 - 2535

25 Venyaminov, S. Y. and Vassilenko, K. S. (1994) Determination of protein tertiary structure class from circular dichroism spectra. Analytical Biochemistry **222**, 176 - 184

26 Sreerame, N. and Woody, R. W. (2000) Estimation of protein secondary structure from circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. Analytical Biochemistry **287**, 252 - 260

27 Sreerame, N. and Woody, R. W. (2004) Computation and analysis of protein circular dichroism spectra. Methods in Enzymology **383**, 318 - 351

28 Poon, S. and Rivers, R. C. (2005) Protein aggregation monitoring on the FLUOstar OPTIMA Microplate reader. http://www.bmglabtech.com/db_assets /applications/ downloads/ applications/ AN135-protein-aggregation-fluorescence-FLUOstar.pdf

29 Gerard, M., Debyser, Z., Desender, L., Kahle, P. J., Baert, J., Baekelandt, V. and Engelborghs, Y. (2006) The aggregation of alpha-synuclein is stimulated by FK506 binding proteins as shown by fluorescence correlation spectroscopy. The FASEB Journal **20**, 524 - 526

30 Sauk, J. J., Johnson, D. and Roszkowski, M. (1982) The effect of some varying lipid A structures on the inhibition of fibrillogenesis in basement membrane collagen. Journal of Oral Pathology **11**, 64 - 71

31 Ecroyd, H. and Carver, J. A. (2008) The effect of small molecules in modulating the chaperone activity of αB-crystallin against ordered and disordered protein aggregation. FEBS Journal **275**, 935 - 947

32 Khurana, R., Coleman, C., Ionescu-Zanetti, C., Carter, S. A., Khrishna, V., Grover, R. K., Roy, R. and Singh, S. (2005) Mechanism of thioflavin T binding to amyloid fibrils. Journal of Structural Biology **151**, 229 - 238

33 Groenning, M., Norrman, M., Flink, J. M., van der Weert, M., Butrinsky, J. T., Schluckebier, G. and Frokjaer, S. (2007) Binding of thioflavin T in insulin amyloid fibrils. Journal of Structural Biology **159**, 483 - 497

34 Saeed, S. M. and Fine, G. (1967) Thioflavin-T for amyloid detection. American Journal of Clinical Pathology **57**, 588 - 593

35 LeVine III, H. (1993) Thioflavine T interaction with synthetic Alzheimer's disease β-amyloid peptides: detection of amyloid aggregation in solution. Protein Science **2**, 404 - 410

36 Krebs, M. R. H., Bromley, E. H. C. and Donald, A. M. (2005) The binding of thioflavin-T to amyloid fibrils: localization and implications. Journal of Structural Biology **149**, 30 - 37

37 Eisert, R., Felau, L. and Brown, L. R. (2006) Methods for enhancing the accuracy and reproducibility of congo red and thioflavin T assays. Analytical Biochemistry **353**, 144 - 146

38 Meehan, S., Berry, Y., Luisi, B., Dobson, C. M., Carver, J. A. and MacPhee, C. E. (2004) Amyloid fibril formation by lens crystallin proteins and its implications for cataract formation. The Journal of Biological Chemistry **279**, 3413 - 3419

39 Jimenez, J. L., Nettleton, E. J., Bouchard, M., Robinson, C. V., Dobson, C. M. and Saibil, H. R. (2002) The protofilament structure of insulin amyloid fibrils. Proceedings of the National Academy of Sciences USA **99**, 9196 - 9201

40 Gross, M., Wilkins, D. K., Pitkeathly, M. C., Chung, E. W., Higham, C., Clark, A. and Dobson, C. M. (1999) Formation of amyloid fibrils by peptides derived from the bacterial cold shock protein CspB. Protein Science **8**, 1350 - 1357

41 Balbirnie, M., Grothe, R. and Eisenberg, D. S. (2001) An amyloid-forming peptide from the yeast prion Sup35 reveals a dehydrated β-sheet structure for amyloid. Proceedings of the National Academy of Sciences USA **98**, 2375 - 2380

42 Makin, O. S. and Serpell, L. C. (2005) Structures for amyloid fibrils. FEBS Journal **272**, 5950 - 5961

43 Serpell, L. C., Fraser, P. E. and Sunde, M. (1999) X-ray fiber diffraction of amyloid fibrils. In Aggregate and Precursor Protein Structure: Aggregate Morphology (Nielsen, E. H., Nybo, M. and Svehag, S.-E., eds.), pp. 526 - 537, Academic Press

44 Bader, R., Bamford, R., Zurdo, J., Luisi, B. F. and Dobson, C. M. (2006) Probing the mechanism of amyloidosis through a tandem repeat of the PI3-SH3 domain suggests a generic model for protein aggregation and fibril formation. Journal of Molecular Biology **356**, 189 - 208

45 Vetri, V., Librizzi, F., Militello, V. and Leone, M. (2007) Effects of succinylation on thermal induced amyloid formation in Concanavalin A. European Biophysical Journal **36**, 733 - 741

46 Schmittschmitt, J. P. and Scholtz, J. M. (2003) The role of protein stability, solubility, and net charge in amyloid fibril formation. Protein Science **12**, 2374 - 2378

47 Veerman, C., de Schiffart, G., Sagis, L. M. C. and van der Linden, E. (2003) Irreversible self-assembly of ovalbumin into fibrils and the resulting network rheology. International Journal of Biological Molecules **33**, 121 - 127

48 Goto, Y., Calciano, L. J. and Fink, A. L. (1990) Acid-induced folding of proteins. Proceedings of the National Academy of Sciences USA **87**, 573 - 577

49 Goto, Y., Takahashi, N. and Fink, A. L. (1990) Mechanism of acid-induced folding of proteins. Biochemistry **29**, 3480 - 3488

50 Kitabatake, N. and Kinekawa, Y.-I. (1995) Turbidity measurement of heated egg proteins using a microplate system. Food Chemistry **54**, 201 - 203

51 Chi, E. Y., Krishnan, S., Randolph, T. W. and Carpenter, J. W. (2003) Physical stability of proteins in aqueous solution: Mechanism and driving forces in nonnative protein aggregation. Pharmaceutical Research **20**, 1325 - 1336

52 Chiti, F., Lorenzi, E. D., Grossi, S., Mangione, P., Giogetti, S., Caccialanza, G., Dobson, C. M., Merlini, G., Ramponi, G. and Bellotti, V. (2001) A partially structured species of β2-microglobulin is significantly populated under physiological conditions and involved in fibrillogenesis. Journal of Biological Chemistry **276**, 16714 - 46721

53 Smith, D. P., Jones, S., Serpell, L. C., Sunde, M. and Radford, S. (2003) A systematic investigation into the effect of protein destabilisation on β2-microglobulin amyloid formation. Journal of Molecular Biology **330**, 943 - 954

54 Wagner, R., Gonzalez, D. H., Podesta, F. E. and Andreo, C. S. (1987) Changes in the quaternary structure of phospho*enol*pyruvate carboxylase induced by ionic strength affect its catalytic activity. European Journal of Biochemistry **164**, 661 - 666

55 Arakawa, T. and Timasheff, S. N. (1984) Mechanism of protein salting in and out by divalent cation salts: Balance between hydration and salt binding. Biochemistry **23**, 5912 - 5923

56 Butcher, S., Hainaut, P. and Milner, J. (1994) Increased salt concentration reversibly destabilizes p53 quaternary structure and sequence-specific DNA binding. Biochemical Journal **298**, 513 - 516

57 Gruninger, R. J., Selinger, L. B. and Mosimann, S. C. (2008) Effect of ionic strength and oxidation on the P-loop conformation of the protein tyrosine phosphatase-like phytase, PhyAsr. FEBS Journal **275**, 3783 - 3792

58 Huang, D.-B., Ainsworth, C. F., Stevens, F. J. and Schiffer, M. (1996) Three quaternary structures for a single protein. Proceedings of the National Academy of Sciences USA **93**, 7017 - 7021

59 Lee, J. Y., Yang, S.-T., Lee, S. K., Jung, H. H., Shin, S. Y., Hahm, K.-S. and Kim, J. I. (2008) Salt-resistant homodimeric bactenecin, a cathelicidin-derived anitmicrobial peptide. FEBS Journal **275**, 3911 - 3920

60 Sidorenko, V. S., Mechetin, G. V., Nevinsky, G. A. and Zharkov, D. O. (2008) Ionic strength and magnesium affect the specificity of *Escherichia coli* and human 8-oxoguanine DNA glycosylases. FEBS Journal **275**, 3747 - 3760

61 Siezen, R. J., Bindels, J. G. and Hoenders, H. J. (1980) The quaternary structure of bovine α-crystallin. European Journal of Biochemistry **111**, 435 - 444

62 Stupak, M., Zoldak, G., Musatov, A., Sprinzl, M. and Sedlak, E. (2006) Unusual effects of salts on the homodimeric structure of NADH oxidase from *Thermus thermophilus* in acidic pH. Biochimica et Biophysica Acta **1764**, 129 - 137

63 Valero, E., De Bonis, S., Filhol, O., Wade, R. H., Langowski, J., Chambaz, E. M. and Cochet, C. (1995) Quaternary structure of casein kinase 2. The Journal of Biological Chemistry **270**, 8345 - 8352

64 Raman, B., Chatani, E., Kihara, M., Ban, T., Sakai, M., Hasegawa, K., Naiki, H., Rao, C. M. and Goto, Y. (2005) Critical balance of electrostatic and hydrophobic interactions is required for β2-microglobulin amyloid fibril growth and stability. Biochemistry **44**, 1288 - 1299

65 Dobson, R. C. J., Valegard, K. and Gerrard, J. A. (2004) The crystal structure of three site-directed mutants of *Escherichia coli* dihydrodipicolinate synthase: Further evidence for a catalytic triad. Journal of Molecular Biology **338**, 329 - 339

66 Dobson, R. C. J., Griffin, M. D. W., Jameson, G. B. and Gerrard, J. A. (2005) The crystal structures of native and (*S*)-lysine-bound dihydrodipicolinate synthase from *Escherichia coli* with improved resolution show new features of biological significance. Acta Crystallographica Section D, Biological Crystallography **D61**, 1116 - 1124

67 Dobson, R. C. J. (2006) DHDPS stability is reduced at -20 ºC over time. pp. Personal communication, Christchurch

68 Pierce (2005) Technical resource: Protein stability and storage.

69 Dobson, R. C. J., Griffin, M. D. W., Roberts, S. J. and Gerrard, J. A. (2004) Dihydrodipicolinate synthase (DHDPS) from *Escherichia coli* displays partial mixed inhibition with respect to its first substrate, pyruvate. Biochimie **86**, 311 - 315

70 Gerrard, J. A. (1992) DPhil thesis. Studies on dihydrodipicolinate synthase, Oxford University, Oxford

71 Mirwaldt, C., Korndorfer, I. and Huber, R. (1995) The crystal structure of dihydrodipicolinate synthase from *Escherichia coli* at 2.5 Å resolution. Journal of Molecular Biology **246**, 227 - 239

72 Andrews, P. (1964) Estimation of the molecular weights of proteins by Sephadex gel filtration. Biochemical Journal **91**, 222 - 233

73 Andrews, P. (1965) The gel filtration behaviour of proteins related to their molecular weights over a wide range. Biochemical Journal **96**, 595 - 606

74 Laber, B., Gomis-Ruth, F., Romao, M. and Huber, R. (1992) *Escherichia coli* dihydrodipicolinate synthase. Identification of the active site and crystallization. Biochemical Journal **288**, 691 - 695

75 Shedlarski, J. G. and Gilvarg, C. (1970) The pyruvate-aspartic semialdehyde condensing enzyme of *Escherichia coli*. The Journal of Biological Chemistry **245**, 1362 - 1373

76 Schuck, P., Perugini, M. A., Gonzales, N. R., Howlett, G. J. and Schubert, D. (2002) Size-distribution analysis of proteins by analytical ultracentrifugation and Lamm equation modeling. Biophysical Journal **78**, 1606 - 1619

77 Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. Biophysical Journal **78**, 1606 - 1619

78 Perugini, M. A., Griffin, M. D. W., Smith, B. J., Webb, L. E., Davin, A. J., Handman, E. and Gerrard, J. A. (2005) Insight into the self-association of key enzymes from pathogenic species. European Biophysical Journal **34**, 469 - 476

79 Yugari, Y. and Gilvarg, C. (1965) The condensation step in diaminopimelate synthesis. The Journal of Biological Chemistry **240**, 4710 - 4716

80 Singh, K. and Bhakuni, V. (2007) Cation induced differential effect on structural and functional properties of *Mycobacterium tuberculosis* α-isopropylmalate synthase. BMC Structural Biology **7**, 39 - 49

81 Manavalan, P. and Johnson Jr., W. C. (1983) Sensitivity of circular dichroism to protein tertiary structure class. Nature **305**, 831 - 832

82 Offerdi, F., Dubail, F., Kischel, P., Sarinski, K., Stern, A. S., Weerdt, C. v. d., Hoch, J. C., Prosperi, C., Francois, J. M., Mayo, S. L. and Martial, J. A. (2003) *De novo* backbone and sequence design of an idealized α/β-barrel protein: Evidence of stable tertiary structure. Journal of Molecular Biology **325**, 163 - 174

83 Unneberg, P., Merelo, J. J., Chacon, P. and Moran, F. (2001) SOMCD: Method for evaluating protein secondary structure from UV circular dichroism spectra. Proteins **42**, 460 - 470

84 Johnson, W. C. (1999) Analyzing protein circular dichroism spectra for accurate secondary structures. Proteins **35**, 307 - 312

85 Pearce, F. G., Perugini, M. A., McKercher, H. J. and Gerrard, J. A. (2006) Dihydrodipicolinate synthase from *Thermotoga maritima*. Biochemical Journal **400**, 359 - 366

86 Yeh, A. P., McMillan, A. and Stowell, M. H. B. (2006) Rapid and simple protein-stability screens: Applications to membrane proteins. Acta Crystallographica Section D: Biological Crystallography **D62**, 451 - 457

87 Leikina, E., Mertts, M. V., Kuznetsova, N. and Leikin, S. (2002) Type I collagen is thermally unstable at body temperature. Proceedings of the National Academy of Sciences USA **99**, 1314 - 1318

88 Jahn, T. R. and Radford, S. (2008) Folding *versus* aggregation: Polypeptide conformations on competing pathways. Archives of Biochemistry and Biophysics. **469**, 100 - 117

89 Clark, P. L. (2004) Protein folding in the cell: Reshaping the folding funnel. Trends in Biochemical Sciences **29**, 527 - 534

90 Schaefer, M., Sommer, M. and Karplus, M. (1997) pH-dependence of protein stability: Absolute electrostatic free energy differences between conformations. Journal of Physical Chemistry **101**, 1663 - 1683

91 Khurana, R., Hate, A. T. and Udgaonkar, J. B. (1995) pH dependence of the stability of Bastar to chemical and thermal denaturation. Protein Science **4**, 1133 - 1144

92 Sabate, R., Gallardo, M. and Estelrich, J. (2003) An autocatalytic reaction as a model for the kinetics of the aggregation of β-amyloid. Biopolymers (Peptide Science) **71**, 190 - 195

93 Stigter, D., Alonso, D. O. V. and Dill, K. A. (1991) Protein stability: Electrostatics and compact denatured states. Proceedings of the National Academy of Sciences USA **88**, 4176 - 4180

94 Hinz, H.-J., Steif, C., Vogel, T., Meyer, R., Renner, M. and Ledermuller, R. (1993) Fundamentals of protein stability. Pure and Applied Chemistry **65**, 947 - 952

95 Wyckoff, R. W. G. (1937) An ultracentrifugal study of the pH stability of tobacco mosaic virus protein. The Journal of Biological Chemistry **117**, 57

96 Nayak, A., Sorci, M., Krueger, S. and Belfort, G. (2008) A universal pathway for amyloid nucleus and precursor formation for insulin. Proteins: Structure, Function, and Bioinformatics **Accepted for publication**

97 Grudzielanek, S., Jansen, R. and Winter, R. (2005) Solvational tuning of the unfolding, aggregation and amyloidogenesis of insulin. Journal of Molecular Biology **351**, 879 - 894

98 Nielsen, L., Khurana, R., Coats, A., Frokjaer, S., Brange, J., Vyas, S., Uversky, V. N. and Fink, A. L. (2001) Effect of environmental factors on the kinetics of insulin fibril formation: Elucidation of the molecular mechanisms. Biochemistry **40**, 6036 - 6046

CHAPTER 3

DO POLYHISTIDINE TAGS INFLUENCE THE PROPERTIES OF *ESCHERCHIA COLI* DHDPS?

3.1 INTRODUCTION

The use of affinity tags to expedite and optimise protein purification is a commonly employed method in modern biochemistry [1-4]. These extra-molecular residues can affect some biophysical characteristics of the protein to which they are attached [2, 5-14].

The impact of a polyhistidine motif on the biophysical properties of the target proteins is sometimes negligible [15-19] but can significantly affect stability, solution properties, oligomeric state, crystallisation and/or binding behavior [2, 5-14]. The effects have also been shown to be dependent on whether the polyhistidine motif is attached to the N- or C-terminus of the protein [8, 10, 20-24]. The potential effects necessitated the extensive characterisation of all the polyhistidine-tagged variants of *E. coli* DHDPS in a systematic fashion [8]. The effect of the affinity tags on the aggregation propensity of *E. coli* DHDPS will be presented in this chapter.

3.2 APPLICATION OF POLYHISTIDINE TAGS TO *E. COLI* DHDPS

The addition of polyhistidine tags to DHDPS was undertaken to expedite and facilitate the production of relatively large quantities of the wild-type protein and its mutants [1-4]. The polyhistidine tags were attached to the N-terminus of DHDPS because reports have indicated that N-terminal polyhistidine tags perform better than those attached to the C-terminus [8, 12, 24]. In addition to this, the N-terminus was some distance away from the active site and was unlikely to interfere with substrate binding or product release [8, 22, 25-27]. Since the mutants, by design, might

be more prone to aggregation, the polyhistidine tags were thought to be necessary to purify sufficient quantities. In addition, the previously employed FPLC techniques (IEX and HIC as per chapter 8, section 8.4.5 and 8.4.6) expose the protein to significant concentrations of NaCl and $(NH_4)_2SO_4$, increasing the risk of aggregation during the purification process itself prior to controlled characterisation.

As discussed in section 3.1, the effect of polyhistidine tags on the properties of DHDPS required investigation. Of particular importance in this comparative study was the potential impact of the polyhistidine tags on the aggregation and amyloidogenic propensity of DHDPS (sections 3.3.3.6 and 3.3.3.7). This was particularly important in the case of the pET M11 polyhistidine tag because previous studies have demonstrated a significant effect of this polyhistidine motif on protein stability [2].

3.3 RESULTS

3.3.1 CHOICE AND LOCATION OF THE POLYHISTIDINE TAGS

Two different polyhistidine tags were used. Both of these constructs (pET M11 DHDPS and pET 151/D-TOPO DHDPS) contained a cleavage site specific for TEV protease; however, the linking regions were variable in length and composition (figure 3.1).

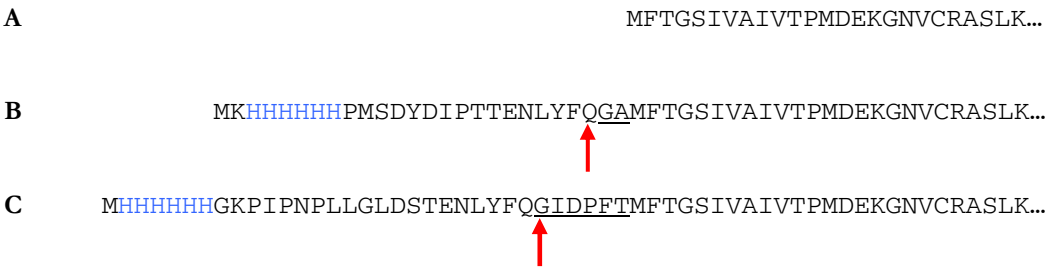


Figure 3.1 – The N-terminal amino acid sequences of **(A)** wild-type DHDPS, **(B)** pET M11 DHDPS and **(C)** pET 151/D-TOPO DHDPS. The residues in blue indicate the polyhistidine motif and the red arrow indicate the cleavage site of TEV protease. The underlined residues represent those remaining following cleavage by TEV protease.

The expression and purification of pET M11 DHDPS was carried out as described in section 3.3.2. The subsequent cleavage of the pET M11 polyhistidine tag was attempted using the protocol described in chapter 8, section 8.4.9. Despite repeated attempts, no cleavage was achieved. This was established by the inability to elute the protein from the affinity columns using a buffer containing a low concentration of imidazole. High imidazole concentrations resulted in elution of the protein and subsequent SDS-PAGE gel electrophoresis confirmed that cleavage was not achieved. From the crystal structure of DHDPS, it was known that the N-terminus was in a shallow pocket on the surface (figure 3.2). As the pET M11 polyhistidine tag was not cleaved upon incubation with the TEV protease, it was hypothesised that the cleavage site was inaccessible to the cleavage enzyme, preventing the cleavage reaction from occurring.

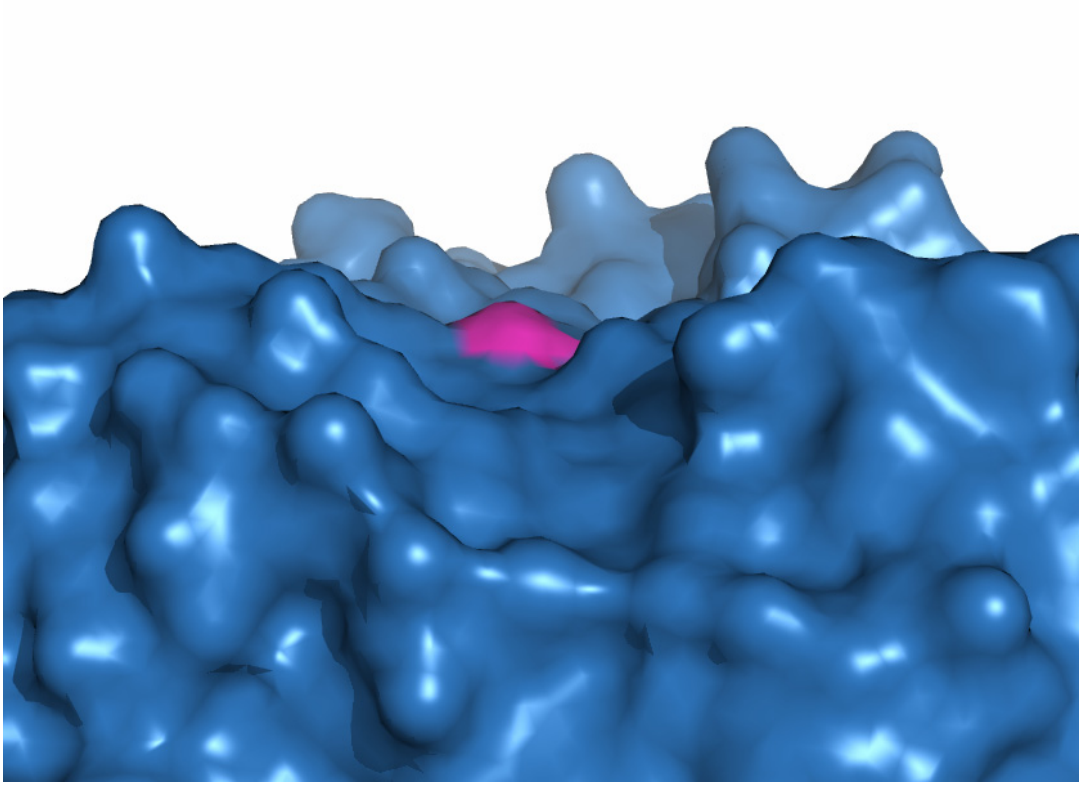


Figure 3.2 – The recess on the surface (blue) of wild-type DHDPS containing the N-terminus (pink) to which the polyhistidine tag was attached. This image was created using PyMol [28] from the *E. coli* DHDPS PDB file (accession number 1yxc) [26].

The second construct (pET 151/D-TOPO DHDPS) also contained a TEV protease cleavage site (figure 3.1). Following purification as described in section 3.3.2, cleavage of this tag was achieved and

confirmed using SDS-PAGE gel electrophoresis. Cleavage was successful, most likely due to the longer linking region between the polyhistidine motif and the N-terminus of DHDPS enabling access of the TEV protease. Six amino acids were left attached to the N-terminus following cleavage (figure 3.1). These residues also had the potential to change the biophysical properties of the protein. Thus both the polyhistidine-tagged and cleaved pET 151/D-TOPO variants of DHDPS were fully characterised and compared to the non-polyhistidine-tagged species.

The extra-molecular residues arising from the addition of the pET M11 polyhistidine tag and the pET 151/D-TOPO polyhistidine tag to DHDPS and the residues remaining following the cleavage of the pET 151/D-TOPO polyhistidine tag, had little effect on the predicted biophysical parameters of the protein. The pI of wild-type DHDPS is 6.44. The pI values for the polyhistidine-tagged variants were similar to the wild-type enzyme (6.57 for pET M11 DHDPS, 6.57 for pET 151/D-TOPO DHDPS and 6.27 for cleaved pET 151/D-TOPO DHDPS) (predicted by "Protein Calculator" (http://www.scripps.edu/~cdputnam/protcalc .html)). The cleaved pET 151/D-TOPO DHDPS had a pI lower than the wild-type enzyme due to the presence of an aspartic acid in the 6 residues remaining following cleavage of the tag.

The aggregation propensities of the polyhistidine-tagged variants of DHDPS were calculated using Zyggregator [29, 30] and compared to the results for the wild-type enzyme. Zyggregator indicated that the aggregation propensities of the polyhistidine-tagged variants of DHDPS were lower than that of wild-type. The Z_{agg} score for wild-type was -4.07. The Z_{agg} scores for pET M11 DHDPS, pET 151/D-TOPO DHDPS and cleaved pET 151/D-TOPO DHDPS were -4.25, -4.26 and -4.15 respectively. Comprehensive results of the Zyggregator analyses are presented in appendix 2, table A2.1.

3.3.2 PURIFICATION

Both tagged and unmodified DHDPS proteins were purified as described in chapter 8, sections 8.4.1 – 8.4.7. The yields for the wild-type DHDPS averaged 19.6 mg of DHDPS (76 units with a specific activity of 3.83 units/mg) per litre of bacterial culture. With the addition of the pET M11 polyhistidine tag, this increased to 30.3 mg/L (130 units with a specific activity of 4.29 units/mg) of bacterial culture. With the addition of the pET 151/D-TOPO polyhistidine tag, 43.4 mg (129 units

with a specific activity of 2.83 units/mg) of protein per litre of bacterial culture was obtained. The yield and purity of the pET M11 DHDPS was improved compared to that of wild-type. The specific activity of the pET 151/D-TOPO DHDPS was lower than wild-type; however, SDS-PAGE gel analysis suggested improved purity (figure 3.3). These results were thought to have been an effect of the pET 151/D-TOPO polyhistidine tag on the activity of the protein. The effect of the tag on kinetics, structure and other biophysical properties was investigated and is presented in the following sections.

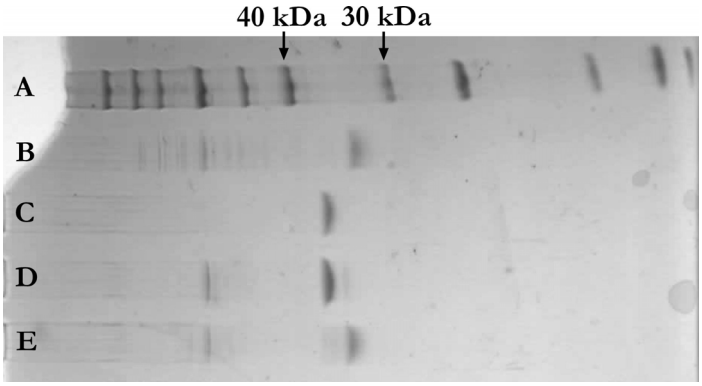


Figure 3.3 – SDS-PAGE. The lanes were loaded with **(A)** molecular weight marker with relevant bands indicated, **(B)** wild-type DHDPS, **(C)** pET M11 DHDPS, **(D)** pET 151-D-TOPO DHDPS and **(E)** cleaved pET 151-D-TOPO DHDPS.

The pET 151/D-TOPO polyhistidine tag was cleaved according to the protocol in chapter 8, section 8.4.9. Some loss of protein was observed during this step due to the incomplete cleavage of the polyhistidine tag from the protein and the subsequent purification steps. The cleavage reaction of the pET 151/D-TOPO DHDPS yielded an average of 67 % of the original protein present. Such losses in yield upon cleavage are typical of affinity tagged oligomeric proteins, arising from the fact that all the monomeric subunits comprising an oligomeric structure need to be cleaved before elution of the protein molecule is possible [13]. Despite this loss of protein, the yield following cleavage averaged 29.1 mg/L (81 units with a specific activity of 2.80 units/mg); a yield comparative to that obtained for wild-type DHDPS. As SDS-PAGE gel electrophoresis revealed that cleaved pET 151/D-TOPO DHDPS contained fewer contaminating proteins than wild-type (figure 3.3), the reduction in specific activity was attributed to the effect of the remaining residues affecting the activity of the protein.

3.3.3 BIOPHYSICAL CHARACTERISATION

Using the suite of techniques described in chapter 2, rigorous biophysical characterisation was undertaken of the polyhistidine-tagged variants of DHDPS. The outcomes of these assays are described in the following sections and the results are compared to those obtained for wild-type DHDPS.

3.3.3.1 KINETICS

The catalytic activities of pET M11 polyhistidine-tagged, pET 151/D-TOPO polyhistidine-tagged and cleaved pET 151/D-TOPO DHDPS were assessed through the measurement of the apparent kinetic parameters using pseudo-single substrate kinetics [31]. According to this analysis, the catalytic capacity of DHDPS increased upon the addition of the pET M11 polyhistidine tag; however, both the pET 151/D-TOPO DHDPS and cleaved pET 151/D-TOPO DHDPS exhibit reduced k_{cat}^{app} for both substrates (table 3.1). It was concluded that the reduction in specific activity (section 3.3.2) observed for pET 151/D-TOPO polyhistidine-tagged and cleaved pET 151/D-TOPO DHDPS was due to the reduction in the catalytic efficiency of the enzyme.

	k_{cat}^{app} with respect to Pyruvate (s^{-1}) ± SEM	**$K_{m\ Pyr}^{app}$ (mM) ± SEM**	**k_{cat}^{app} with respect to ((*S*)-ASA (s^{-1}) ± SEM**	**$K_{m\ (S)\text{-}ASA}^{app}$ (mM) + SEM**
Wild-type DHDPS	118 ± 3	0.3 ± 0.0	123 ± 5	0.1 ± 0.0
pET M11 DHDPS	127 ± 2	0.1 ± 0.0	142 ± 7	0.3 ± 0.0
pET 151/D-TOPO DHDPS	85 ± 2	0.3 ± 0.0	92 ± 6	0.8 ± 0.1
Cleaved pET 151/D-TOPO DHDPS	94 ± 2	0.1 ± 0.0	81 ± 4	0.4 ± 0.1

Table 3.1 – The kinetic parameters for the polyhistidine-tagged variants of DHDPS compared to those of wild-type. The k_{cat}^{app} values for all the polyhistidine-tagged variants are within the range of published values (k_{cat} ranging from 79 to 188 ± 7 s^{-1} [27, 32, 33]) as are the $K_{m\ Pyr}^{app}$ (K_{mPyr} ranging from 0.1 ± 0.01 mM to 0.26 ± 0.03 mM [27, 32, 33]). The $K_{m\ (S)\text{-}ASA}^{app}$ for all polyhistidine-tagged variants are higher than that of the wild-type ($K_{m\ (S)\text{-}ASA}$ ranging from 0.11 ± 0.01 mM to 0.13 ± 0.02 mM) [27, 32, 33]. The two k_{cat} values arise from the use of pseudo-single substrate models rather than a complete two substrate analysis. The errors are the SEM from > 3 replicates.

The apparent Michaelis constant for (*S*)-ASA was increased for all variants whilst the K_m^{app} for pyruvate is the same or less than wild-type. These effects were attributed to the presence of the polyhistidine tags and the residues remaining following cleavage of the pET 151/D-TOPO polyhistidine tag. It is possible that the presence of the extra-molecular residues reduced conformational stability of the protein, altering the kinetic parameters, as reported for dimeric DHDPS [32] as well as other proteins [34, 35].

3.3.3.2 CRYSTALLOGRAPHIC ANALYSIS

An X-ray crystal structure of DHDPS was sought to investigate the impact of the tag on protein structure. As a positive control, wild-type DHDPS was crystallised in parallel, using the technique described in Mirwaldt *et al.* (1995) [36] as modified by Dobson *et al.* (2005) [26]. The wild-type crystals were grown at 12 ºC, appeared after approximately 3 days and formed a trigonal shape approximately 0.3 mm across, consistent with previous reports. Repeated attempts were made to crystallise pET M11 DHDPS and pET 151/D-TOPO DHDPS under the same conditions and at 4 ºC; however, these attempts were unsuccessful. In addition to the standard conditions, the polyhistidine-tagged variants were also subjected to sparse-matrix screens (Hampton Research screens HR2-110 and HR2-112) in an effort to find favorable crystallisation conditions. These yielded no promising results; however, upon removal of the pET 151/D-TOPO polyhistidine tag, crystals formed under the standard conditions, forming trigonal structures similar to those formed by wild-type DHDPS. The structure was subsequently solved by Pearce (unpublished data). The crystals of cleaved pET 151/D-TOPO DHDPS obtained by Pearce diffracted to 1.9 Å (unpublished data). Alignment of this structure with the crystal structure obtained by Dobson *et al.* (2005) [26] (PDB accession number 1yxc) showed no significant difference in the secondary, tertiary or quaternary fold of the protein (figure 3.4). The amino acids remaining following cleavage of the pET 151/D-TOPO polyhistidine tag were not visible in the crystal structure due to a high level of disorder. The inherent disorder of this structure provides a clue to the inability of the polyhistidine-tagged variants to form crystals; if the polyhistidine tags were disordered, then crystallisation may have been inhibited [5].

As the polyhistidine tags increased the size of DHDPS it is likely that the crystallisation conditions were significantly altered. As alternative methods were available for the characterisation of the polyhistidine-tagged variants, crystallographic analyses of the polyhistidine variants were not pursued.

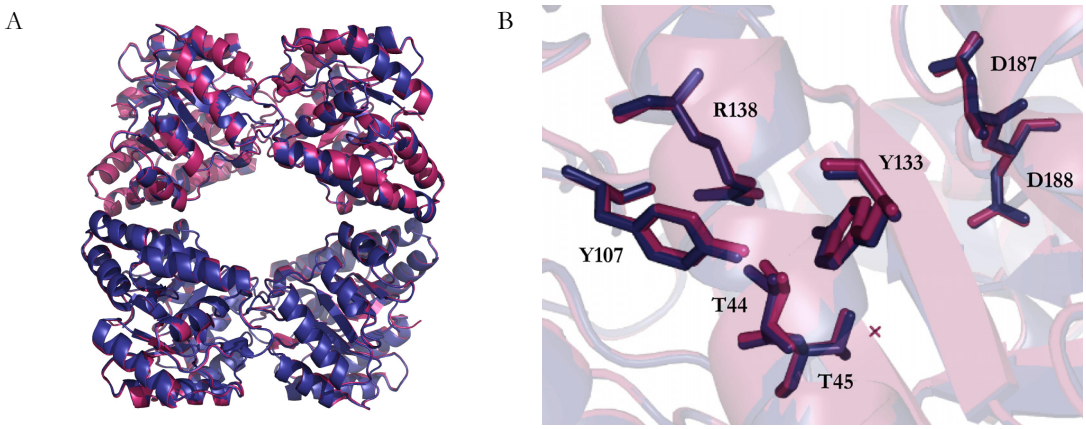


Figure 3.4 – **(A)** Alignment of the structure of wild-type DHDPS (blue) (Dobson *et al.* [26]) with the structure of cleaved pET 151/D-TOPO DHDPS (pink) (provided by Pearce) demonstrated no significant alterations to either the secondary, tertiary or quaternary structures. The rmsd for the tetramer = 0.268 Å. **(B)** Superposition of the residues involved in the active site exhibited no significant structural changes in the crystalline state. This image was created using PyMol [28] from the *E. coli* DHDPS PDB file (accession number 1yxc) [26] and the crystal structure of cleaved pET 151/D-TOPO DHDPS.

The fact that the polyhistidine tags altered the ability of DHDPS to form crystals suggested there may have been a shift in the equilibrium between conformational states (chapter 1, figure 1.1), a hypothesis supported by the changed kinetic parameters (section 3.3.3.1). As well as reducing the propensity towards crystallisation, this shift could have resulted in proteins with an increased predisposition to form amorphous or β-sheet-specific aggregates. The altered properties may also have affected the ability of the protein to associate into its native homotetrameric conformation. Thus testing of the quaternary structures of pET M11 DHDPS, pET 151/D-TOPO DHDPS and cleaved pET 151/D-TOPO DHDPS in solution was undertaken.

3.3.3.3 ANALYTICAL GEL PERMEATION LIQUID CHROMATOGRAPHY

Analytical gel permeation liquid chromatography was used to assess the quaternary structures of pET M11 DHDPS, pET 151/D-TOPO DHDPS and cleaved pET 151/D-TOPO DHDPS in solution.

The results were compared to that of wild-type DHDPS. Figure 3.5 shows the chromatograms obtained for each of the DHDPS variants compared to that of wild-type DHDPS.

The molecular masses of wild-type DHDPS, pET M11 DHDPS, pET 151/D-TOPO DHDPS and cleaved pET 151/D-TOPO DHDPS were calculated from their V_e and the calibration plot as described in chapter 2, section 2.3.2.3. The molecular masses were all within the 10 % error of the expected molecular masses from analytical gel permeation liquid chromatography data [37, 38] and also within 10 % error of the range of molecular masses previously established using gel filtration (112 – 126 kDa) [39-41]. The results confirm that all the variants exist as tetramers in solution. These results indicate that attaching a polyhistidine tag had no substantial affect on the quaternary structure of DHDPS, thus analytical ultra-centrifugation was not performed on these variants.

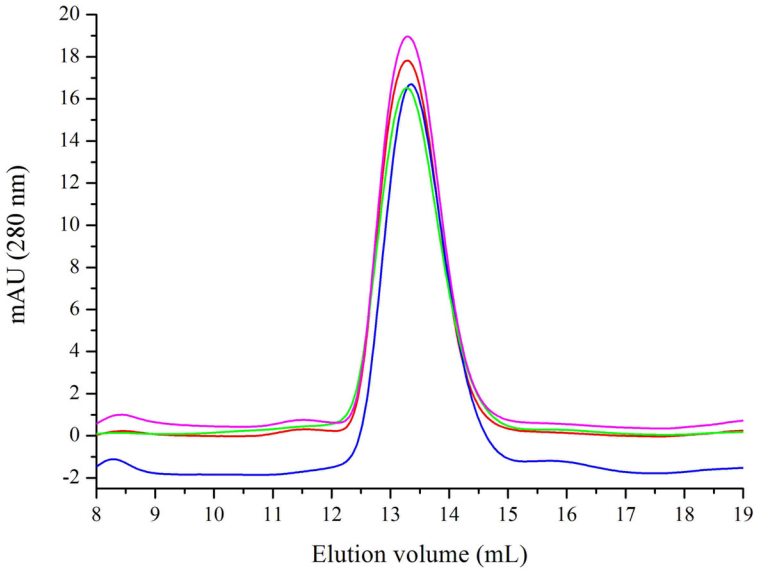


Figure 3.5 – Elution profiles determined by zonal analysis for wild-type DHDPS (red), pET M11 DHDPS (green), pET151/D-TOPO DHDPS (blue) and cleaved pET151/D-TOPO DHDPS (pink) all exhibited elution volumes of 13.3 mL (in 20 mM Tris (pH 7.0), 100 mM NaCl). This indicated that all the variants existed as tetramers. The chromatograms were produced by plotting the absorbance at 280 nm as a function of elution volume (V_e).

3.3.3.4 DIFFERENTIAL SCANNING FLUORIMETRY

The thermal denaturation temperatures of pET M11 DHDPS, pET 151/D-TOPO DHDPS and cleaved pET 151/D-TOPO DHDPS were assessed using a thermofluor technique as described in

chapter 2, section 2.1.5. The results are shown in figure 3.6, compared to those obtained for wild-type DHDPS. Statistical analyses of the results were carried out as described in chapter 2, section 2.3.2.5. The analysis of deviance table for all DHDPS variants is presented in appendix 1, table A1.2.

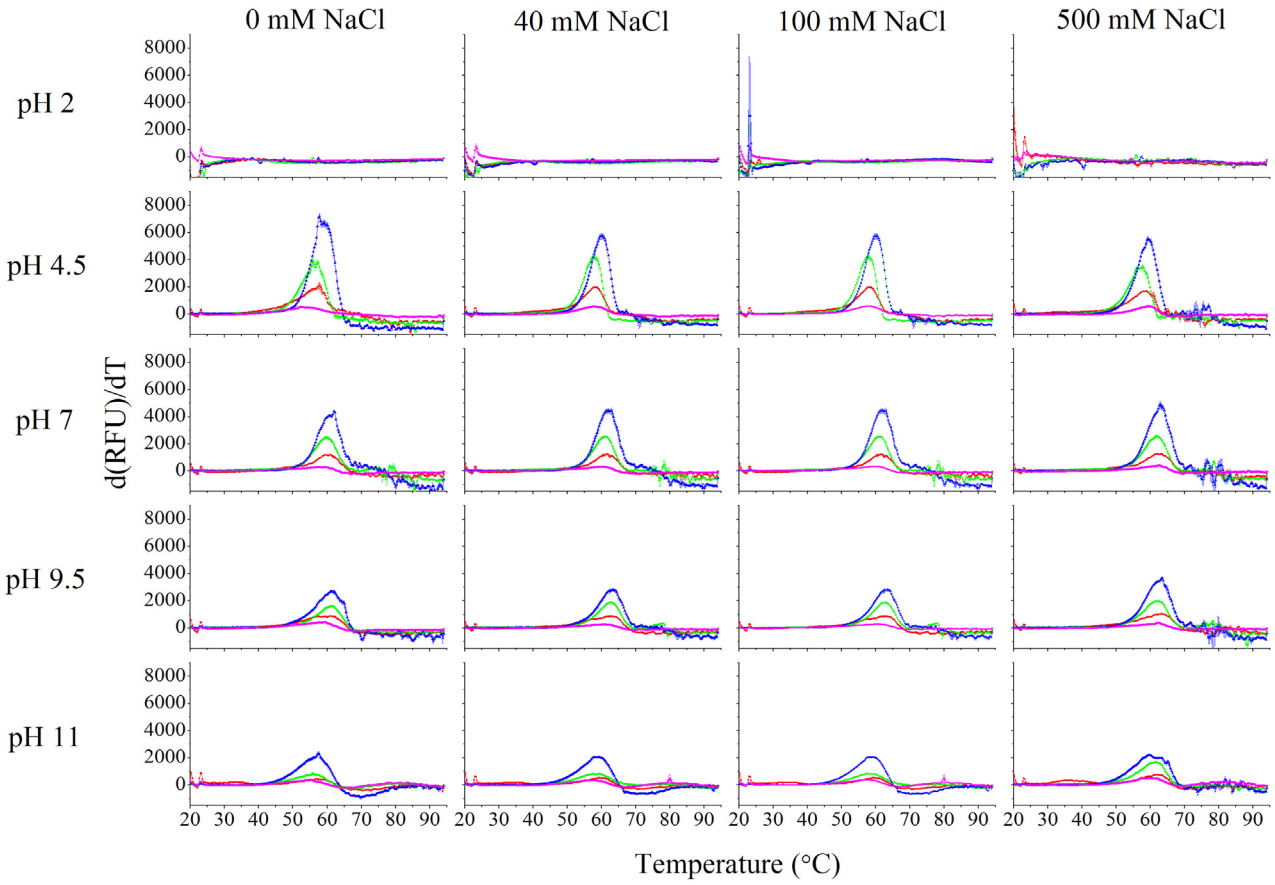


Figure 3.6 – The thermal denaturation profiles of wild-type DHDPS (red), pET M11 DHDPS (green), pET 151/D-TOPO DHDPS (blue) and cleaved pET 151/D-TOPO DHDPS (pink) at 0.5 mg/mL in 100 mM phosphate buffer (pH 7, 40 mM NaCl) as monitored by SYPRO orange fluorescence. The data plotted are the derivatives of the increase in fluorescence monitored by the BioRad IQ5 (presented in appendix A2, figure A2.1). The peaks indicate the temperature at which the protein is unfolding the most rapidly and is recorded as the mean thermal denaturation temperature. The data plotted are the mean values of three replicates ± SEM.

The results of this assay indicated that the presence of a either a pET M11 or pET 151/D-TOPO polyhistidine tag or the amino acids that remained after the cleavage of the pET 151/D-TOPO polyhistidine tag did not significantly affect the thermal denaturation temperature of DHDPS; however, there was a non-significant increase in thermal stability for some proteins under some

conditions. The underlying causes of inconsistent values obtained for the maximum RFU are discussed in chapter 2, section 2.3.2.5. As these causes were unlikely to result in a change in the thermal denaturation temperature, no further investigation was carried out into this phenomenon [42].

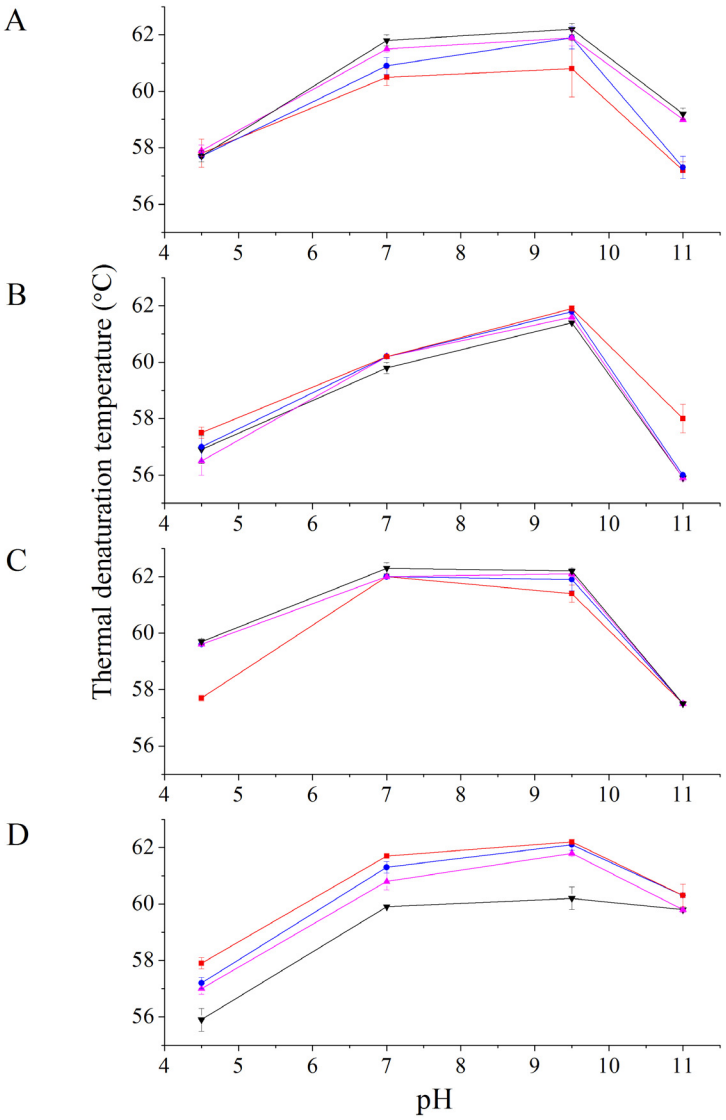


Figure 3.7 – The thermal denaturation temperatures of **(A)** wild-type DHDPS, **(B)** pET M11 DHDPS, **(C)** pET 151/D-TOPO DHDPS and **(D)** cleaved pET 151/D-TOPO DHDPS at 0.5 mg/mL in 100 mM phosphate buffer at 0 mM NaCl (red), 40 mM NaCl (blue), 100 mM NaCl (pink) and 500 mM NaCl (black) plotted as a function of pH. The mean thermal denaturation temperatures plotted were calculated from derivations of the fluorescence data. The data plotted are the mean values of three replicates and the error bars represent the SEM.

As described for wild-type DHDPS, the effect of salt on the thermal denaturation temperatures of the polyhistidine-tagged variants appeared to be small; however, the effect of pH on the thermal denaturation of the polyhistidine-tagged variants, like wild-type, was significant (see figure 3.7). At pH 2, all the polyhistidine-tagged variants exhibited high initial fluorescence suggesting that hydrophobic patches were exposed prior to the addition of the dye [43]. As discussed in chapter 2, section 2.3.2.5, the high initial fluorescence was unlikely be due to complete denaturation of the protein. At pH 7 and pH 9.5 the proteins were the most thermally stable. The pI values for the polyhistidine-tagged variants were similar to the wild-type enzyme (6.44 for wild-type DHDPS, 6.57 for pET M11 DHDPS, 6.57 for pET 151/D-TOPO DHDPS and 6.27 for cleaved pET 151/D-TOPO DHDPS). Thus the stability at pH 7 and pH 9.5 was not surprising, arising from the deprotonation of the proteins causing an increase in solubility and therefore thermal stability [44, 45]. At pH 11 and 4.5 the thermal denaturation temperature of pET M11 DHDPS was significantly reduced, as observed for the wild-type enzyme. The effects of pH 4.5 and pH 11 on pET 151/D-TOPO DHDPS and cleaved pET 151/D-TOPO DHDPS were not so easily interpreted. At pH 4.5 the thermal stability of pET 151/D-TOPO DHDPS was increased, although not significantly. At pH 11 the thermal stability of cleaved pET 151/D-TOPO DHDPS was increased significantly to a temperature close to that observed for pH 7 and pH 9.5. These results did not seem to be related to the pI of the proteins.

DSF demonstrated that the presence of a polyhistidine tag caused little change to the thermal denaturation of DHDPS under a range of conditions. However, some proteins exhibited increased thermal stability at high and low pH. This increased stability was further investigated by characterising the amorphous and β-sheet-specific aggregation as described in the following sections (sections 3.3.3.6 and 3.3.3.7). The results of the DFS were corroborated using CD spectroscopy.

3.3.3.5 CIRCULAR DICHROISM SPECTROSCOPY

The secondary structures of pET M11 DHDPS, pET 151/D-TOPO DHDPS and cleaved pET 151/D-TOPO DHDPS were assessed using circular dichroism (CD) spectroscopy and compared to the results obtained for wild-type DHDPS. Figure 3.8 shows that the wavelength spectra of the variants were very similar. This assay demonstrated that the presence of the polyhistidine tags had no effect on the secondary structure of DHDPS in solution. All the spectra had double minima at 208

nm and 222 nm which are the characteristic of $(\alpha/\beta)_8$ barrels [46-48] and consistent with the spectra previously obtained for wild-type DHDPS [32, 49, 50].

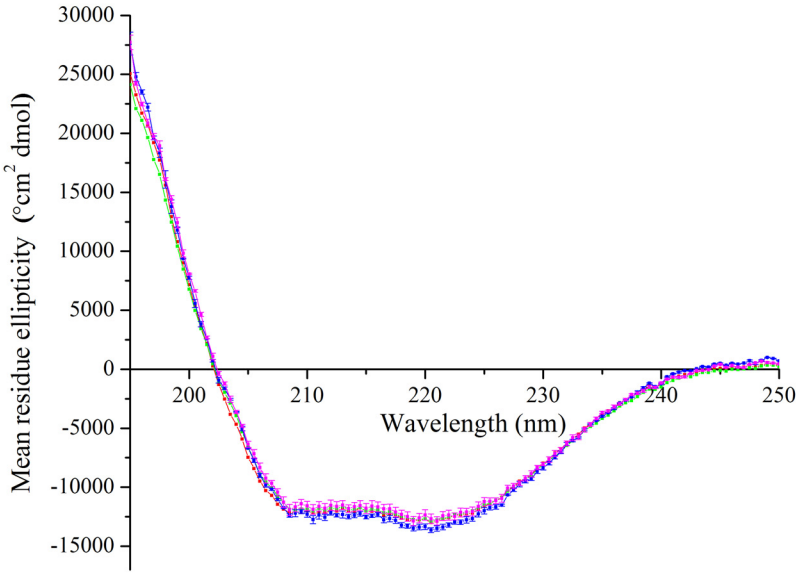


Figure 3.8 – Wavelength scans of wild-type DHDPS (red), pET M11 DHDPS (green), pET 151/D-TOPO DHDPS (blue) and cleaved pET 151/D-TOPO DHDPS (pink) at 0.2 mg/mL in 20 mM phosphate buffer (pH 7) containing 40 mM NaCl at 0.2 mg/mL. The plot shows mean residue ellipticity vs. wavelength ± SEM (N = 3).

Thermal melts of pET M11 DHDPS, pET 151/D-TOPO DHDPS and cleaved pET 151/D-TOPO DHDPS were performed by monitoring the change in intensity at 220 nm. All demonstrated similar denaturation profiles to that of wild-type DHDPS (see table 3.2 and appendix 2, section A2.3).

Variant	Thermal denaturation temperature as per CD (°C)	Thermal denaturation temperature as per DSF (°C)
Wild-type DHDPS	56.9	60.9 ± 0.3
pET M11 DHDPS	57.0	60.2 ± 0.0
pET 151/D-TOPO DHDPS	61.2	62.0 ± 0.0
Cleaved pET 151/D-TOPO DHDPS	58.6	60.8 ± 0.3

Table 3.2 – The temperatures at which wild-type DHDPS, pET M11 DHDPS, pET 151/D-TOPO DHDPS and cleaved pET 151/D-TOPO DHDPS melted in 20 mM phosphate buffer (pH 7, 40 mM NaCl) as monitored by CD spectroscopy and DSF. The errors on the DSF data represent the SEM of three replicates. There are no error bars on the CD data due to the high level of reproducibility as discussed in chapter 2, section 2.3.2.6.

By CD spectroscopy, the secondary structure of pET M11 DHDPS, pET 151/D-TOPO DHDPS and cleaved pET 151/D-TOPO DHDPS appeared very similar to wild-type. The thermal denaturation temperatures as determined by the CD demonstrated that pET 151/D-TOPO DHDPS and cleaved pET 151/D-TOPO DHDPS melted at significantly higher temperatures than the wild-type enzyme. This increased stability correlated with the wider range of stability demonstrated by DSF. The variability observed between the CD and DSF techniques for all the polyhistidine-tagged variants is small, consistent with the literature on other proteins [42, 51, 52]. The results of both techniques were qualitatively similar, indicating a thermal denaturation temperature of approximately 60 ºC in pH 7, 40 mM NaCl.

3.3.3.6 AMORPHOUS AGGREGATION

The amorphous aggregation propensities of pET M11 DHDPS, pET 151/D-TOPO DHDPS and cleaved pET 151/D-TOPO DHDPS were measured across all chosen conditions (see chapter 2, section 2.2). The results obtained for each variant were compared to the results for wild-type DHDPS. Each assay was carried out as described in chapter 2, section 2.3.2.7 and chapter 8, section 8.6.5.

The results of this assay suggested that the presence of a polyhistidine tag significantly affected the aggregation potential of DHDPS, generally increasing the rate of aggregation whilst reducing the maximum absorbance observed. The effect of the presence of the polyhistidine tags and the 6 residues remaining following the cleavage of the pET 151/D-TOPO polyhistidine tag on the aggregation of DHDPS in 100 mM phosphate buffer (pH 7, 40 mM NaCl) is shown in figure 3.9. The effects of the presence of polyhistidine tags on the DHDPS on amorphous aggregation under all conditions are summarised in appendix 2, section A2.4.

Where possible, sigmoid (Boltzmann) functions were fitted to the aggregation data. As discussed in chapter 2 section 2.3.2.7 the sigmoid functions fitted to the data allowed an assessment of the aggregation half life ($Agg_{½}$). Some data were not able to be analysed in this way due to extremely rapid aggregation or no increase in light scattering. Maximum fluorescence was calculated from the data gathered during the initial 90 minutes of the assay. The results are summarised in figure 3.9 and presented in their entirety in appendix 2, table A2.3. Statistical analysis was carried out using a multifactor ANOVA, with protein species, pH, and salt as categorical factors. Significance is

presented to a P value of < 0.01. The analysis of deviance for $Agg_{½}$ is presented in appendix 2, table A2.4. Only the results at pH 7 and pH 9.5 were used in the statistical analysis as the data sets at other pH were incomplete due to rapid aggregation or no measurable absorbance. The analysis of deviance for maximum absorbance is presented in appendix 2, table A2.5.

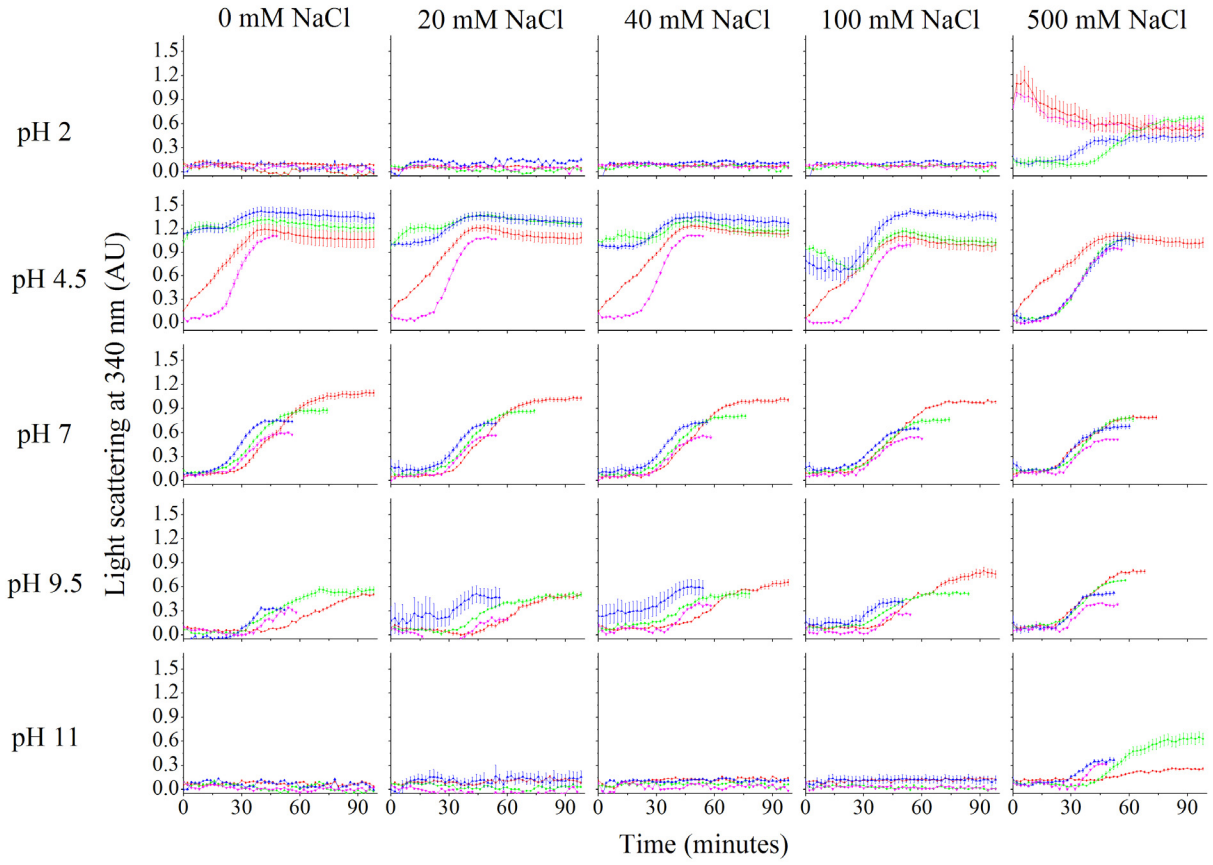


Figure 3.9 – The aggregation profiles of wild-type DHDPS (red), pET M11 DHDPS (green), pET 151/D-TOPO DHDPS (blue) and cleaved pET 151/D-TOPO DHDPS (pink) in 100 mM phosphate buffer at 60ºC. Where possible, a sigmoid (Boltzmann) function was used for the line of best fit (> 98% confidence interval) in order to calculate the aggregation half life ($Agg_{½}$). The $Agg_{½}$ was determined by identifying the midpoint of the transition as previously described [53]. In order to fit the sigmoid curve, only the data up until the linear phase were included. Once the peak absorbance was reached the light scattering was inconsistent due to the sedimentation of the protein aggregate. The maximum turbidity was the highest reading over the first 90 minutes of the assay. The data plotted are the mean values of six replicates and the error bars represent the SEM.

The effect of salt concentration on the polyhistidine-tagged variants was different to that observed for wild-type DHDPS. The results of the assay demonstrated that most salt concentrations had little

effect on either the $Agg_{½}$ or the maximum light scattering for any of the polyhistidine-tagged variants. The proteins in buffer containing 500 mM NaCl exhibited the most variability in their propensity to form amorphous aggregate (particularly at pH 2 and pH 11). The reduction in solvating power of NaCl [44] in combination with the extra-molecular residues may have led to the variability observed. As it is widely accepted that particulate size affects light scattering [54], the differences in maximum absorbance (between conditions and DHDPS variants) were attributed variable particulate size.

The effect of pH on the aggregation of the DHDPS variants was significant. All variants exhibited the highest maximum light scattering at pH 4.5 followed by pH 7 and pH 9.5. Only very low readings were obtained for pH 2 and pH 11, except for in the presence of 500 mM NaCl, which appeared to cause an increase in aggregation. The aggregation half life at pH 4.5 was too rapid to quantify; however, the $Agg_{½}$ values for pH 7 and pH 9.5 were calculated and compared. The speed of aggregation was not significantly different between pH 7 and pH 9 for any of the polyhistidine-tagged species. This differed from the results obtained for wild-type which underwent more rapid aggregation at pH 9.5 than at pH 7.

Comparison of the aggregation propensities of the protein revealed that both pET M11 and pET 151/D-TOPO polyhistidine-tagged variants underwent more rapid aggregation than the wild-type protein, resulting in higher maximum light scattering at pH 4.5 (0 – 100 mM NaCl) and pH 11, 500 mM NaCl but reduced maximum light scattering at pH 7 and pH 9.5. The rate of aggregation of the cleaved pET 151/D-TOPO DHDPS was generally slower than that of wild-type and demonstrated a lower maximum light scattering. At pH 4.5, 500 mM NaCl all variants aggregated more slowly than wild-type, suggesting that the salt had a stabilising affect on the polyhistidine-tagged variants. The observed variability at pH 4.5 was unlikely to be a function of the altered isoelectric points of the polyhistidine-tagged species, as they were not dissimilar to that of wild-type (see section 3.3.3.4)

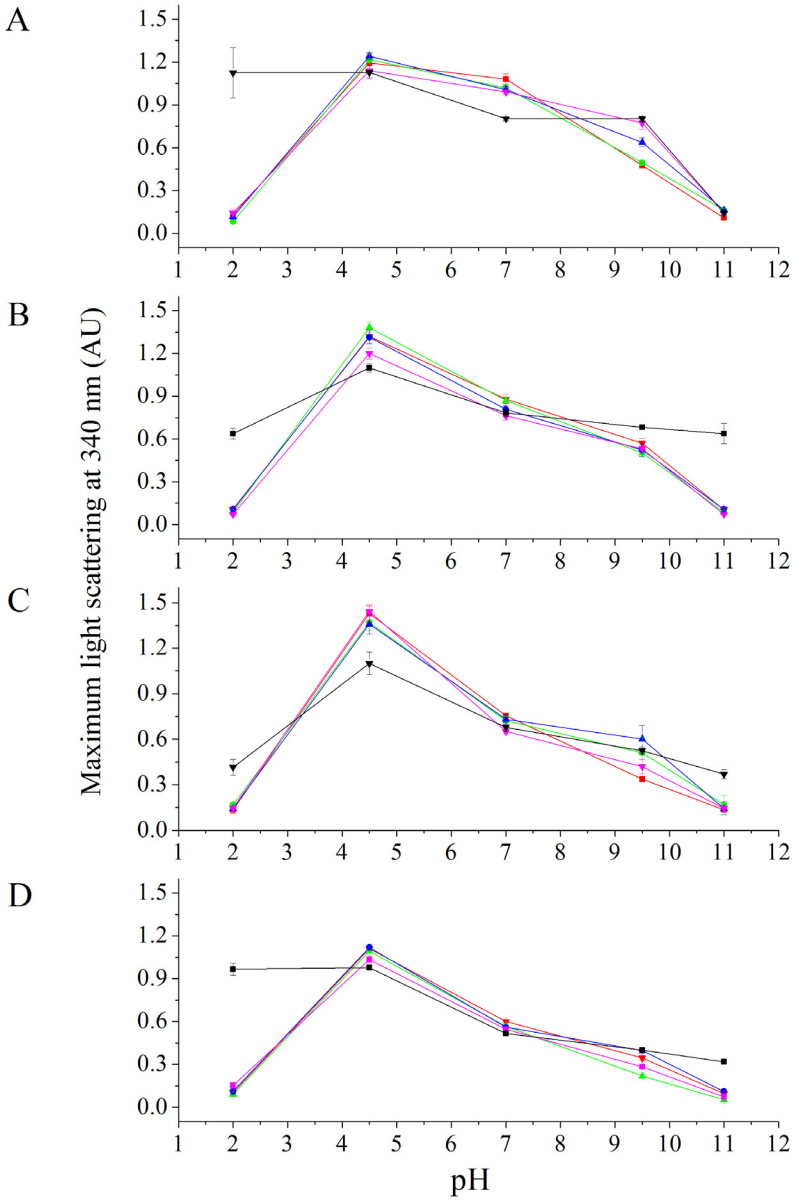


Figure 3.10 – The maximum light scattering of **(A)** wild-type DHDPS, **(B)** pET M11 DHDPS, **(C)** pET 151/D-TOPO DHDPS and **(D)** cleaved pET 151/D-TOPO DHDPS in 100 mM phosphate buffer 0 mM NaCl (red), 20 mM NaCl (green), 40 mM NaCl (blue), 100 mM NaCl (pink) and 500 mM NaCl (black). The data plotted are the mean maximum absorbance units reached over 90 minutes (n = 6) and are plotted as a function of pH. The error bars represent the SEM.

The results of this assay indicated that the presence of pET M11 polyhistidine tag, pET 151/D-TOPO polyhistidine tag and the residues remaining following cleavage of the pET 151/D-TOPO polyhistidine tag, significantly affected the propensity of DHDPS to aggregate. Along with altered propensities to form amorphous aggregates, the extra residues potentially altered the capacity for β-

sheet rich aggregate formation by DHDPS. Thus ThT assays were performed to quantify the β-sheet formation and are described in section 3.3.3.7.

3.3.3.7 β-SHEET-SPECIFIC AGGREGATION

The results of these assays suggested that the presence of the polyhistidine tags or amino acids remaining following the cleavage of the pET 151/D-TOPO polyhistidine tag, do significantly affect the speed of β-sheet-specific aggregation of DHDPS. These assays were carried out as described in chapter 2 section 2.3.2.8 and chapter 8, section 8.6.6 and the results were compared to those obtained for wild-type DHDPS.

The effects of the polyhistidine tags on DHDPS with regard to ThT fluorescence (in 100 mM phosphate buffer is shown in figure 3.11, and are summarised in figure 3.12 and appendix 2, section A2.5. Statistical analysis was carried out using a multifactor ANOVA with protein as the categorical factor. The data residuals for the maximum fluorescence were subjected to statistical analysis using a multifactor ANOVA with protein species, salt and pH as the categorical factors. pH and salt were included in the model because they have an effect on ThT fluorescence however the effect is not relevant due to their effect on ThT binding. The main effect of protein is the independent effect after controlling between the differences across pH and salt concentrations. Significance is presented to a P value of < 0.01. The analysis of deviance for $\beta\text{-}agg_{½}$ is presented in appendix 2, table A2.7. Only the results at pH 4.5 and pH 7 were used in the statistical analysis of the $\beta\text{-}agg_{½}$ as the data sets at other pHs were incomplete due to rapid aggregation or no measurable absorbance. The analysis of deviance for maximum absorbance is presented in appendix 2, table A2.8.

The ThT fluorescence in the presence of the polyhistidine-tagged variants of DHDPS were compared to the results obtained for the wild-type enzyme and showed that the addition of the polyhistidine motifs and the residues remaining following the cleavage of the pET 151/D-TOPO polyhistidine tag significantly altered the maximum fluorescence. The pET M11 DHDPS demonstrated significantly higher maximum fluorescence than the wild-type protein under all salt conditions (except at pH 4.5 in the presence of 500 mM NaCl). In this case, the increase was not significant. pET 151/D-TOPO DHDPS incubated with ThT exhibited significantly higher maximum fluorescence than wild-type under all conditions. The only exceptions to this were at pH

4.5, 500 mM NaCl and pH 7, 500 mM NaCl, at which the maximum fluorescence was similar wild-type. The cleaved pET 151/D-TOPO DHDPS, upon incubation with ThT, demonstrated a significantly lower maximum fluorescence than that of wild-type. The $\beta\text{-}agg_{½}$ did not significantly change for any variant, suggesting that the rate of β-sheet-specific aggregation played no role in the maximum fluorescence observed. As discussed in chapter 2, section 2.3.2.8, neither the effects of pH or salt on the β-sheet-specific aggregation can be compared due to their effects on ThT fluorescence.

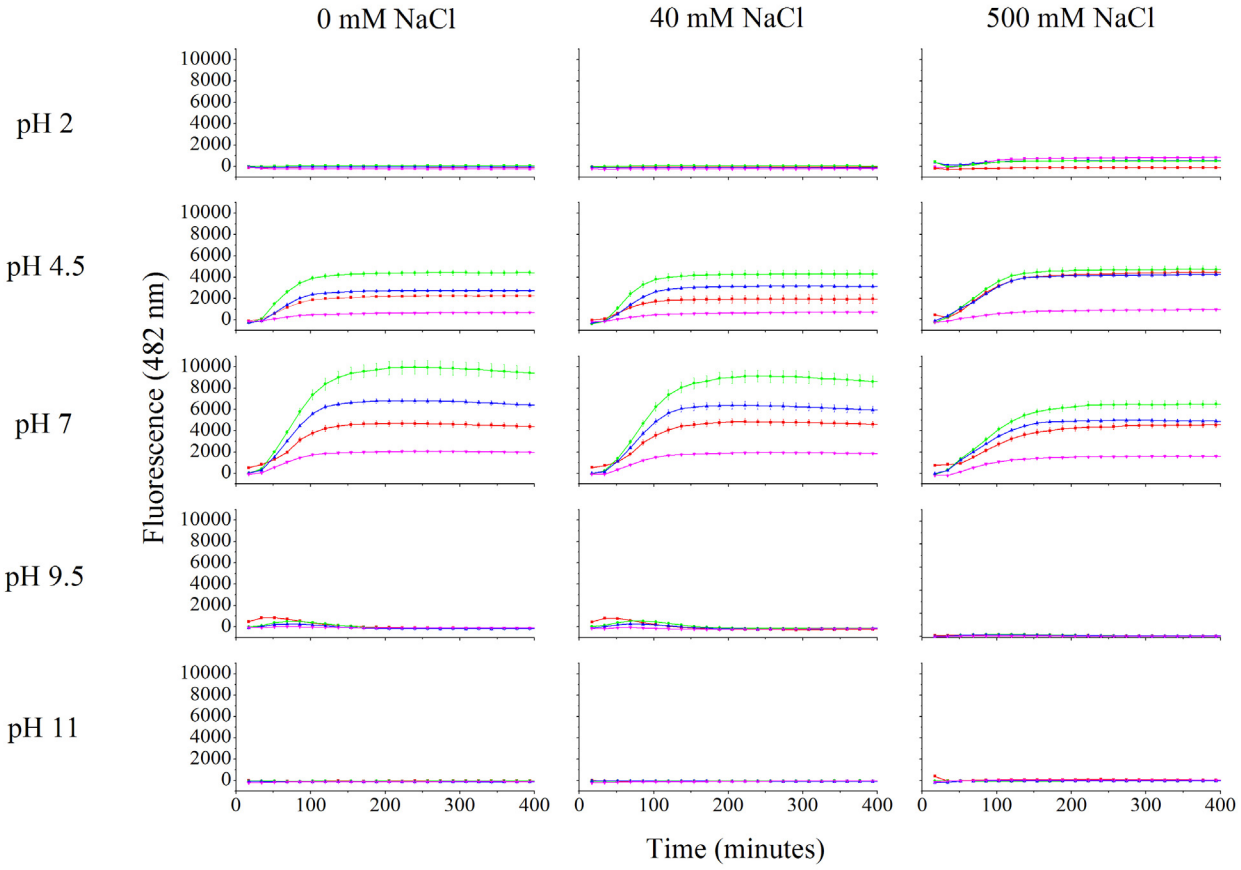


Figure 3.11 – The ThT fluorescence in the presence of wild-type DHDPS (red), pET M11 DHDPS (green), pET 151/D-TOPO DHDPS (blue) and cleaved pET 151/D-TOPO DHDPS (pink) in 100 mM phosphate buffer at 60°C. The data plotted are the mean values of six replicates ± SEM. Where possible, a sigmoid (Boltzmann) function was used for the line of best fit (> 98% confidence interval) in order to calculate the β-sheet-specific aggregation half life ($\beta\text{-}agg_{½}$). The $\beta\text{-}agg_{½}$ was determined by identifying the midpoint of the transition as previously described [53]. Note that only the data up until the linear phase was included in the analysis of the $\beta\text{-}agg_{½}$ [55]. Once the maximum fluorescence was reached the fluorescence readings were inconsistent due to sedimentation of the protein aggregate [55]. The maximum fluorescence was the highest reading over the initial 300 minutes of the assay.

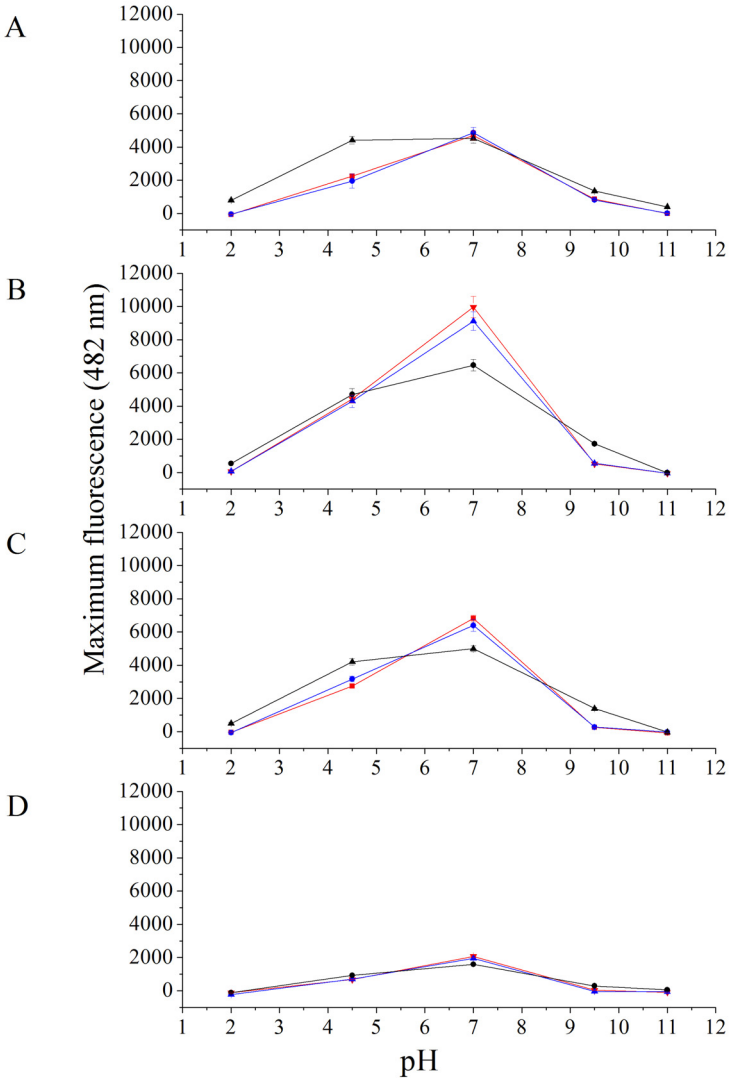


Figure 3.12 – The aggregation of **(A)** wild-type DHDPS, **(B)** pET M11 DHDPS, **(C)** pET 151/D-TOPO DHDPS and **(D)** cleaved pET 151/D-TOPO DHDPS in 0 mM NaCl (red), 40 mM NaCl (blue), and 500 mM NaCl (black). The data plotted are the mean maximum fluorescence reached over 300 minutes and are plotted as a function of pH ± SEM (n=6).

The effects of the presence of polyhistidine tags on the propensity of DHDPS to form β-sheet-specific aggregate were significant. The addition of the pET M11 polyhistidine tag increased the total fluorescence observed in the presence of ThT. The pET 151/DTOPO polyhistidine tag also increased the maximum fluorescence observed compared to wild-type DHDPS; however, upon cleavage, the propensity to form β-sheet rich aggregate was reduced to less than that of wild-type.

The significant increase in ThT fluorescence observed suggested that both the pET M11 DHDPS and pET 151/D-TOPO DHDPS form more β-sheet-specific aggregate than the wild-type protein. As fluorescence of ThT in the presence of β-sheet-specific aggregate is considered indicative of amyloid formation [56-60], attempts to confirm the presence of fibrillar structures were pursued through both TEM and X-ray fibre diffraction.

3.3.3.8 CONFIRMATION OF AMYLOID FIBRILS

Confirmation of the presence of amyloid fibrils by TEM was attempted. Samples containing pET M11 DHDPS and pET 151/D-TOPO DHDPS were prepared under the same conditions as those that exhibited the highest ThT induced fluorescence (described in chapter 8, section 8.8). Despite repeated efforts, no fibrillar structures were observed (figure 3.13). Cleaved pET 151/D-TOPO DHDPS was not assayed in this manner as the ThT fluorescence was consistently lower than that of the wild-type enzyme, for which no fibrils were observed. In addition to this, attempts to construct X-ray fibre diffraction stalks were unsuccessful (described in chapter 8, section 8.8.2). The failure to confirm the presence of amyloid fibrils suggests that the ThT results were not indicative of classical amyloid formation. This was consistent with the fact that $(\alpha/\beta)_8$ barrels have not been shown to form amyloid fibrils and that polyhistidine tags have not been reported to increase the propensities of proteins to form amyloid fibrils.

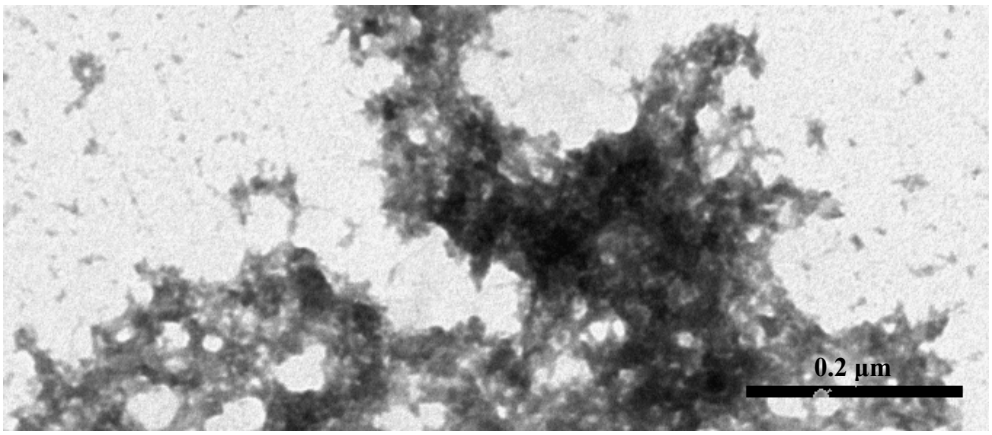


Figure 3.13 – TEM image of non-specific aggregate formed by pET M11 DHDPS in 100 mM phosphate buffer (pH 7) 40 mM NaCl, at 60 ºC. This image was typical of all 5 images obtained and characteristic of the images obtained for pET 151/D-TOPO polyhistidine-tagged *E. coli* DHDPS.

3.4 SUMMARY AND CONCLUSIONS

The addition of polyhistidine tags to *E. coli* DHDPS affected the enzyme's kinetic properties, causing a reduction in the k_{cat}^{app} and an increase in the K_m^{app} for (*S*)-ASA. Crystallography showed that the cleaved pET 151/D-TOPO polyhistidine-tagged variant had similar secondary, tertiary and quaternary structure to that of the wild-type protein. Analytical gel permeation liquid chromatography confirmed these results, indicating that all the polyhistidine-tagged variants existed as tetramers in solution. CD spectra showed that the secondary structures of each of the polyhistidine-tagged variants were similar to that of the wild-type enzyme. The thermal stabilities of the polyhistidine-tagged species were similar to that of wild-type DHDPS as determined by DSF and CD spectroscopy.

An analysis of the propensity of pET M11 DHDPS, pET 151/D-TOPO DHDPS and cleaved pET 151/D-TOPO DHDPS to form amorphous and β-sheet-specific aggregate was carried out. The presence of the polyhistidine motifs increased the maximum absorbance and the $Agg_{½}$ of DHDPS in some conditions. Such destabilisation has previously been reported for proteins containing polyhistidine constructs [2, 5-14]. The more surprising results were obtained from the analysis of β-sheet-specific aggregation propensity of the polyhistidine-tagged species. This demonstrated that the presence of both the pET M11 and pET 151/D-TOPO polyhistidine tags significantly increased the maximum ThT fluorescence. This suggests the polyhistidine motif leads to an elevated tendency for the formation of β-sheet rich aggregates and therefore potentially amyloid fibrils, although, in this case, the presence of fibrils was not confirmed. This has significant implications for any application of polyhistidine motifs as well as other extra-molecular sequences to proteins. Upon addition of these sequences, the predisposition of proteins to form aggregates may increase and as environment plays an important role in determining the capacity to aggregate, the exposure of these constructs to new conditions via use in technology or industry may result in unexpected and unpredictable results. This was highlighted by discrepancies between the results of Zyggregator and the empirical data. As Zyggregator is designed to predict β-sheet-specific aggregation from an unfolded state, the differences are not surprising. Despite the aromatic nature of the histidine residues, Zyggregator does not consider them to be conducive to promoting β-sheet-specific aggregation. This assumption was based on research by Azreil *et al.* who associated all hydrophobic amino acid except histidines with amyloidogenic sequences [61]. The results described in this chapter suggest that the presence of

the polyhistidine motif may increase the propensity to form amorphous and possible β-sheet rich aggregates.

This comparative control study demonstrated the impact of the polyhistidine tags on the amorphous and β-sheet-specific aggregation propensity of DHDPS (sections 3.3.3.6 and 3.3.3.7). Because the tags had a significant effect on a range of the biophysical parameters of the enzyme, the mutations affecting the tertiary and quaternary structure of the protein detailed in the following chapters were investigated with and without the polyhistidine tags in a combinatorial fashion.

3.5 REFERENCES

1 Sørensen, H. P. and Mortensen, K. K. (2005) Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. Journal of Biotechnology **115**, 113 - 128

2 Dümmler, A., Lawrence, A.-M. and de Marco, A. (2005) Simplified screening for the detection of soluble fusion constructs expressed in *E. coli* using a modular set of vectors. Microbial Cell Factories **13**, 34

3 Arnau, J., Lauritzen, C., Petersen, G. E. and Pedersen, J. (2006) Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. Protein Expression and Purification **48**, 1 - 13

4 Kimple, M. E. and Sondek, J. (2004) Overview of affinity tags for protein purification. Current Protocols in Protein Science **Supplement 36**, 9.9.1 - 9.9.19

5 Bucher, M. H., Evdokimov, A. G. and Waugh, D. S. (2002) Differential effects of short affinity tags on the crystallization of *Pyrococcus furiosis* maltodextrin-binding protein. Acta Crystallographica Section D: Biological Crystallography **D58**, 392 - 397

6 Rumlová, M., Benedíková, J., Cubínková, R., Pichová, I. and Ruml, T. (2001) Comparison of classical and affinity purification techniques of Mason-Pfizer monkey virus capsid protein: The alteration of the product by affinity tag. Protein Expression and Purification **23**, 75 - 83

7 Dolja, V. V., Peremyslov, V. V., Keller, K. E., Martin, R. R. and Hong, J. (1998) Isolation and stability of histidine-tagged proteins produced in plants via potvirus gene vectors. Virology **252**, 269 - 274

8 Goel, A., Colcher, D., Koo, J.-S., Booth, B. J. M., Pavlinkova, G. and Batra, S. K. (2000) Relative position of the hexahistidine tag effects binding properties of a tumor-associated single-chain Fv construct. Biochimica et Biophysica Acta **1523**, 13 - 20

9 Hammarström, M., Hellgren, N., van den Berg, S., Berglund, H. and Härd, T. (2002) Rapid screening for improved solubility of small human proteins produced as fusion proteins in *Escherichia coli*. Protein Science **11**, 313 - 321

10 Mohanty, A. K. and Wiener, M. C. (2004) Membrane protein expression and production: effects of polyhistidine tag length and position. Protein Expression and Purification **33**, 311 - 325

11 Routzahn, K. M. and Waugh, D. S. (2002) Differential effects of supplementary affinity tags on the solubility of MBP fusion proteins. Journal of Structural and Functional Genomics **2**, 83 - 92

12 Woestenenk, E. A., Hammarström, M., van den Berg, S., Härd, T. and Berglund, H. (2004) His tag effect on solubility of human proteins produced in *Escherichia coli*: A comparison between four expression vectors. Journal of Structural and Functional Genomics **5**, 217 - 229

13 Kenig, M., Peternel, S., Gaberc-Porekar, V. and Menart, V. (2006) Influence of the protein oligomericity on final yield after affinity tag removal in purification of recombinant proteins. Journal of Chromatography A **1101**, 293 - 306

14 Smyth, D. R., Mrokiewicz, M. K., McGrath, W. J., Listwan, P. and Kobe, B. (2003) Crystal structures of fusion proteins with large-affinity tags. Protein Science **12**, 1313 - 1322

15 Waldner, J. C., Lahr, S. J., Edgell, M. H. and Pielak, G. J. (1998) Effect of a polyhistidine terminal extension on eglin c stability. Analytical Biochemistry **263**, 116 - 118

16 Reid, K. L., Rodriquez, H. M., Hillier, B. J. and Gregoret, L. M. (1998) Stability and folding properties of a model β-sheet protein, *Escherichia coli* CspA. Protein Science **7**, 470 - 479

17 Li, Z. and Crooke, E. (1999) Functional analysis of affinity-purified polyhistidine-tagged DnaA protein. Protein Expression and Purification **17**, 41 - 48

18 Hou, E. W., Prasad, R., Beard, W. A. and Wilson, S. H. (2004) High-level expression and purification of untagged and histidine-tagged HIV-1 reverse transcriptase. Protein Expression and Purification **34**, 75 - 86

19 Chatterjee, D. K. and Esposito, D. (2006) Enhanced soluble protein expression using two new fusion tags. Protein Expression and Purification **46**, 122 - 129

20 Xu, C.-G., Fan, X.-J., Fu, Y.-J. and Liang, A.-H. (2008) Effect of location of the His-tag on the production of soluble and functional *Buthus martensii* Karsch insect toxin. Protein Expression and Purification **59**, 103 - 109

21 Hammarström, M., Woestenenk, E. A., Hellgren, N., Härd, T. and Berglund, H. (2006) Effect of N-terminal solubility enhancing fusion proteins on yield of purified target protein. Journal of Structural and Functional Genomics **7**, 1 - 14

22 Mason, A. B., He, Q.-Y., Halbrooks, P. J., Everse, S. J., Gumerov, D. R., Kaltashov, I. A., Smith, V. C., Hewitt, J. and MacGillivray, R. T. A. (2002) Differential effect of a his tag at the N- and C-termini: Functional studies with recombinant human serum transferrin. Biochemistry **41**, 9448 - 9454

23 Sachdev, D. and Chirgwin, J. M. (1998) Order of fusions between bacterial and mammalian proteins can determine solubility. Biochemical and Biophysical Research Communication **244**, 933 - 937

24 Busso, D., Kim, R. and Kim, S.-H. (2003) Expression of soluble recombinant proteins in a cell-free system using a 96-well format. Journal of Biochemical and Biophysical Methods **55**, 233 - 240

25 Dobson, R. C. J., Gerrard, J. A. and Pearce, F. G. (2004) Dihydrodipicolinate synthase is not inhibited by its substrate, (*S*)-aspartate β-semialdehyde. Biochemical Journal **377**, 757 - 762

26 Dobson, R. C. J., Griffin, M. D. W., Jameson, G. B. and Gerrard, J. A. (2005) The crystal structures of native and (*S*)-lysine-bound dihydrodipicolinate synthase from *Escherichia coli* with improved resolution show new features of biological significance. Acta Crystallographica Section D, Biological Crystallography **D61**, 1116 - 1124

27 Dobson, R. C. J., Griffin, M. D. W., Roberts, S. J. and Gerrard, J. A. (2004) Dihydrodipicolinate synthase (DHDPS) from *Escherichia coli* displays partial mixed inhibition with respect to its first substrate, pyruvate. Biochimie **86**, 311 - 315

28 DeLano, W. L. (2002) The PyMOL molecular graphics system. DeLano Scientific, San Carlos

29 Pawar, A. P., DuBay, K. F., Zurdo, J., Chiti, F., Vendruscolo, M. and Dobson, C. M. (2005) Prediction of "aggregation-prone" and "aggregation-susceptible" regions in proteins associated with neurodegenerative diseases. Journal of Molecular Biology **350**, 379-392

30 DuBay, K. F., Pawar, A. P., Chiti, F., Zurdo, J., Dobson, C. M. and Vendruscolo, M. (2004) Prediction of absolute aggregation rates of amyloidogenic polypeptide chains. Journal of Molecular Biology **341**, 1317 - 1326

31 Cornish-Bowden, A. (1999) Fundamentals of enzyme kinetics. Portland Press Ltd., London

32 Griffin, M. D., Dobson, R. C. J., Pearce, F. G., Antonio, L., Whitten, A. E., Liew, C. K., Mackay, J. P., Trewhella, J., Jameson, G. B., Perugini, M. A. and Gerrard, J. A. (2008) Evolution of quaternary structure in a homotetrameric enzyme. Journal of Molecular Biology **380**, 691 - 703

33 Karsten, W. E. (1997) Dihydrodipicolinate synthase for *Escherichia coli*: pH dependent changes in the kinetic mechanism and kinetic mechanism of allosteric inhibition by *L*-lysine. Biochemistry **36**, 1730 - 1739

34 Borchert, T. V., Kishan, K. V., Zeelan, J. P. and al, e. (1995) Three new crystal structures of point mutation variants of monoTIM: Conformational flexibility of loop-1, loop-4 and loop-8. Structure **3**, 669 - 679

35 Griffin, M. D. and Gerrard, J. A. (2008) The relationship between oligomeric state and protein function. In Protein Dimerization (and Oligomerization) in Biology. (Matthews, J., ed.), Landes Bioscience, Sydney

36 Mirwaldt, C., Korndorfer, I. and Huber, R. (1995) The crystal structure of dihydrodipicolinate synthase from *Escherichia coli* at 2.5 Å resolution. Journal of Molecular Biology **246**, 227 - 239

37 Andrews, P. (1964) Estimation of the molecular weights of proteins by sephadex gel filtration. Biochemical Journal **91**, 222 - 233

38 Andrews, P. (1965) The gel filtration behaviour of proteins related to their molecular weights over a wide range. Biochemical Journal **96**, 595 - 606

39 Griffin, M. D. W. (2005) PhD Thesis. Why is DHDPS a tetramer?, University of Canterbury, Christchurch

40 Laber, B., Gomis-Ruth, F., Romao, M. and Huber, R. (1992) *Escherichia coli* dihydrodipicolinate synthase. Identification of the active site and crystallization. Biochemical Journal **288**, 691 - 695

41 Shedlarski, J. G. and Gilvarg, C. (1970) The pyruvate-aspartic semialdehyde condensing enzyme of *Escherichia coli*. The Journal of Biological Chemistry **245**, 1362 - 1373

42 Epps, D. E., Sarver, R. W., Rogers, J. M., Herberg, J. T. and Tomich, P. K. (2001) The ligand affinity of proteins measured by isothermal denaturation kinetics. Analytical Biochemistry **292**, 40 - 50

43 Vedadi, M., Niesen, F. H., Allali-Hassani, A., Fedorov, O. Y., Finerty, P. J. J., Wasney, G. A., Yeung, R., Arrowsmith, C., Ball, L. J., Berglund, H., Hui, R., Marsden, B. D., Norlund, P., Sundstrom, M., Weigelt, J. and Edwards, A. M. (2006) Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. Proceedings of the National Academy of Sciences USA **103**, 15835 - 15840

44 Arakawa, T. and Timasheff, S. N. (1984) Mechanism of protein salting in and out by divalent cation salts: Balance between hydration and salt binding. Biochemistry **23**, 5912 - 5923

45 Chi, E. Y., Krishnan, S., Randolph, T. W. and Carpenter, J. W. (2003) Physical stability of proteins in aqueous solution: Mechanism and driving forces in nonnative protein aggregation. Pharmaceutical Research **20**, 1325 - 1336

46 Singh, K. and Bhakuni, V. (2007) Cation induced differential effect on structural and functional properties of *Mycobacterium tuberculosis* α-isopropylmalate synthase. BMC Structural Biology **7**, 39 - 49

47 Manavalan, P. and Johnson Jr., W. C. (1983) Sensitivity of circular dichroism to protein tertiary structure class. Nature **305**, 831 - 832

48 Offerdi, F., Dubail, F., Kischel, P., Sarinski, K., Stern, A. S., Weerdt, C. v. d., Hoch, J. C., Prosperi, C., Francois, J. M., Mayo, S. L. and Martial, J. A. (2003) *De novo* backbone and sequence design of an idealized α/β-barrel protein: Evidence of stable tertiary structure. Journal of Molecular Biology **325**, 163 - 174

49 Pearce, F. G., Dobson, R. C., Weber, A., Lane, L. A., McCammon, M. G., Squire, M. A., Perugini, M. A., Jameson, G. B., Robinson, C. V. and Gerrard, J. A. (2008) Mutating the tight-dimer interface of dihydrodipicolinate synthase disrupts the enzyme quaternary structure: towards a monomeric enzyme. Biochemistry **In press**

50 Dobson, R. C. J., Devenish, S. R. A., Turner, L. A., Clifford, V. R., Pearce, F. G., Jameson, G. B. and Gerrard, J. A. (2005) Role of arginine 138 in the catalysis and regulation of *Escherichia coli* dihydrodipicolinate synthase. Biochemistry **44**, 13007 - 13013

51 Yeh, A. P., McMillan, A. and Stowell, M. H. B. (2006) Rapid and simple protein-stability screens: Applications to membrane proteins. Acta Crystallographica Section D: Biological Crystallography **D62**, 451 - 457

52 Leikina, E., Mertts, M. V., Kuznetsova, N. and Leikin, S. (2002) Type I collagen is thermally unstable at body temperature. Proceedings of the National Academy of Sciences USA **99**, 1314 - 1318

53 Sabate, R., Gallardo, M. and Estelrich, J. (2003) An autocatalytic reaction as a model for the kinetics of the aggregation of β-amyloid. Biopolymers (Peptide Science) **71**, 190 - 195

54 Nairn, J. A. (2003) Polymer characterisation. Materials Science and Engineering **54**, 89 - 100

55 Niesen, F. H., Berglund, H. and Vedadi, M. (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. Nature Protocols **2**, 2212 - 2221

56 Khurana, R., Coleman, C., Ionescu-Zanetti, C., Carter, S. A., Khrishna, V., Grover, R. K., Roy, R. and Singh, S. (2005) Mechanism of thioflavin T binding to amyloid fibrils. Journal of Structural Biology **151**, 229 - 238

57 Groenning, M., Norrman, M., Flink, J. M., van der Weert, M., Butrinsky, J. T., Schluckebier, G. and Frokjaer, S. (2007) Binding of thioflavin T in insulin amyloid fibrils. Journal of Structural Biology **159**, 483 - 497

58 Saeed, S. M. and Fine, G. (1967) Thioflavin-T for amyloid detection. American Journal of Clinical Pathology **57**, 588 - 593

59 LeVine III, H. (1993) Thioflavine T interaction with synthetic Alzheimer's disease β-amyloid peptides: Detection of amyloid aggregation in solution. Protein Science **2**, 404 - 410

60 Krebs, M. R. H., Bromley, E. H. C. and Donald, A. M. (2005) The binding of thioflavin T to amyloid fibrils: localization and implications. Journal of Structural Biology **149**, 30 - 37

61 Azriel, R. and Gazit, E. (2001) Analysis of the minimal amyloid-forming fragment of the islet amyloid polypeptide - An experimental support for the key role of phenylalanine residue in amyloid formation. Journal of Biological Chemistry **276**, 34156 - 34161

CHAPTER 4

TESTING ALGORITHMIC PREDICTION

4.1 INTRODUCTION

As discussed in chapter 1, there have been a number of algorithms designed to predict the aggregation propensity of proteins. Several of these algorithms can be used to predict the effect of amino acid substitutions on protein stability. This can aid the optimisation of protein expression through the production of variants with increased solubility [1-6]. The accuracies of the algorithms have been tested against sets of proteins with known propensities to aggregate [1-6]. The reported accuracies are high; however, testing the algorithms with regard to proteins with uncharacterised aggregation potentials provides a more objective test.

DHDPS was used as a model for assessing the variability between the predictions of two algorithms, Zyggregator and TANGO (see section 4.2). DHDPS is an $(\alpha/\beta)_8$ barrel, a class of proteins that (as discussed in chapter 1) has never been shown to form amyloid fibrils. The substitution of amino acids that increase the propensity of DHDPS to form amyloid fibrils (as predicted by Zyggregator) provides an opportunity, not only to test the accuracy of the algorithm but also to potentially create an $(\alpha/\beta)_8$ barrel protein with a tendency to form amyloid fibrils.

4.2 ALGORITHMIC PREDICTION OF *E. COLI* DHDPS AGGREGATION PROPENSITY

Zyggregator

The overall aggregation propensity of the DHDPS amino acid chain was calculated using Zyggregator. The conditions were set at the standard values (pH 7, ionic strength of 0.05 M, protein concentration 0.034 mg/mL). The results of the analysis indicated that DHDPS has a relatively low

propensity to aggregate, as determined by the Z_{agg} score (where a Z_{agg} score of < 0 is less prone to aggregate than randomly generated sequences of the same length under the same conditions) [2]. As a point of comparison for the Z_{agg} results, the published score for peptide sequences derived from the protein implicated in Alzheimer's disease (Aβ) range up to 2.52 and scores for peptides derived from α-synuclein (the protein involved in Parkinson's disease) range up to 1.89 [2].

Overall hydrophobicity	95.38
Overall charge	-5
Overall β-sheet propensity	1090.06
Overall α-helical propensity	1133.27
Number of hydrophobic/hydrophilic patterns	2
Absolute amyloid aggregation rate (k)	-4.07128
Intrinsic aggregation propensity	-6.12471
Z_{agg} score	-0.458258

Table 4.1 – The results of the prediction of the overall aggregation propensity of DHDPS. The Z_{agg} score is a function of the other parameters provided. The overall hydrophobicity, overall charge, overall β-sheet propensity, and overall α-helical propensity are presented for the protein as a whole and are dependent on the amino acid composition. These parameters, combined with the number of hydrophobic and hydrophilic patterns are used by the algorithm to predict the absolute aggregation rate, the aggregation propensity and therefore the Z_{agg} score. As the Z_{agg} score is less than 0, the propensity of DHDPS to aggregate is relatively low.

Zyggregator can also be used to identify amino acids that promote the conversion of the protein as a whole into an amyloidogenic species [2]. This prediction is based on the biophysical properties of the amino acids and assigns a value to each based on how likely they are to increase the likelihood of amyloid fibril formation. Generally, the hydrophobic residues have a higher tendency to increase the aggregation propensity of a protein as a whole. The only exception to this is histidine, which according to Azriel *et al.* [7], occurs at a much lower frequency in amyloidogenic peptides than other aromatic amino acids, thus it was assigned a much lower amyloid aggregation propensity than other aromatic residues. The summation of the individual intrinsic factors for each amino acid in a polypeptide sequence resulted in an "aggregation propensity profile" (figure 4.1). This considered the wild-type sequence and the effect of all possible amino acid replacements at each position, enabling

the identification of regions that have a potential to increase or decrease the overall aggregation propensity of the protein.

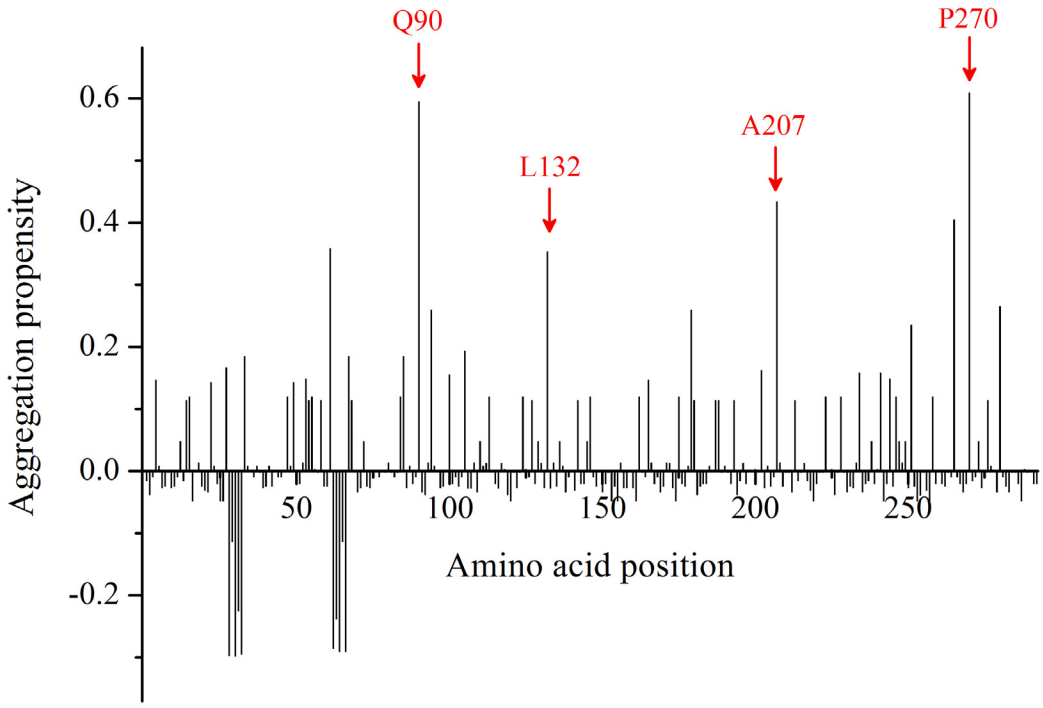


Figure 4.1 – The "aggregation propensity profile" for DHDPS. The peaks indicated are those residues with potential to, upon amino acid substitution, increase aggregation propensity of the protein as a whole.

TANGO

TANGO was used to predict the cross-β aggregation propensity (*agg)* of DHDPS at the settings comparable with those used for Zyggregator (pH 7, 25 ºC, and at an ionic strength of 0.05 M). The results of the analysis are summarised in table 4.2 along with the predictions concerning the propensity of DHDPS to form α-helical aggregates (*helagg*), turns (*turn)* and α-helix (*helix).*

Like Zyggregator, TANGO can be used to estimate an "aggregation propensity profile" of the protein (see figure 4.2). TANGO was used to gauge the relative aggregation sensitivity of an index in the sequence, 19 amino acids were substituted in place of the wild type amino acid and a (*agg*) number generated. The 20 TANGO runs for each index were averaged. The difference between the wild-type aggregation propensity and average mutated aggregation propensity for a given amino acid

in the sequence implies a qualitative difference in aggregation propensity between the wild-type sequence and points mutants at that position. This identifies the amino acids in the polypeptide, that upon substitution cause an increase in the β-aggregation propensity of the protein as a whole.

Cross-β aggregation (*agg*)	704.22
α-helical aggregation (*helagg*)	2.09 x 10 $^{-07}$
Turns (*turn*)	65.08
α-helices (*helix)*	78.49

Table 4.2 - The results of TANGO with regard to wild-type DHDPS. The algorithm considers that every segment of protein can occupy any state according to a Boltzmann distribution with the frequency of each structural state being proportional to its energy which is estimated from statistical and empirical considerations [1].

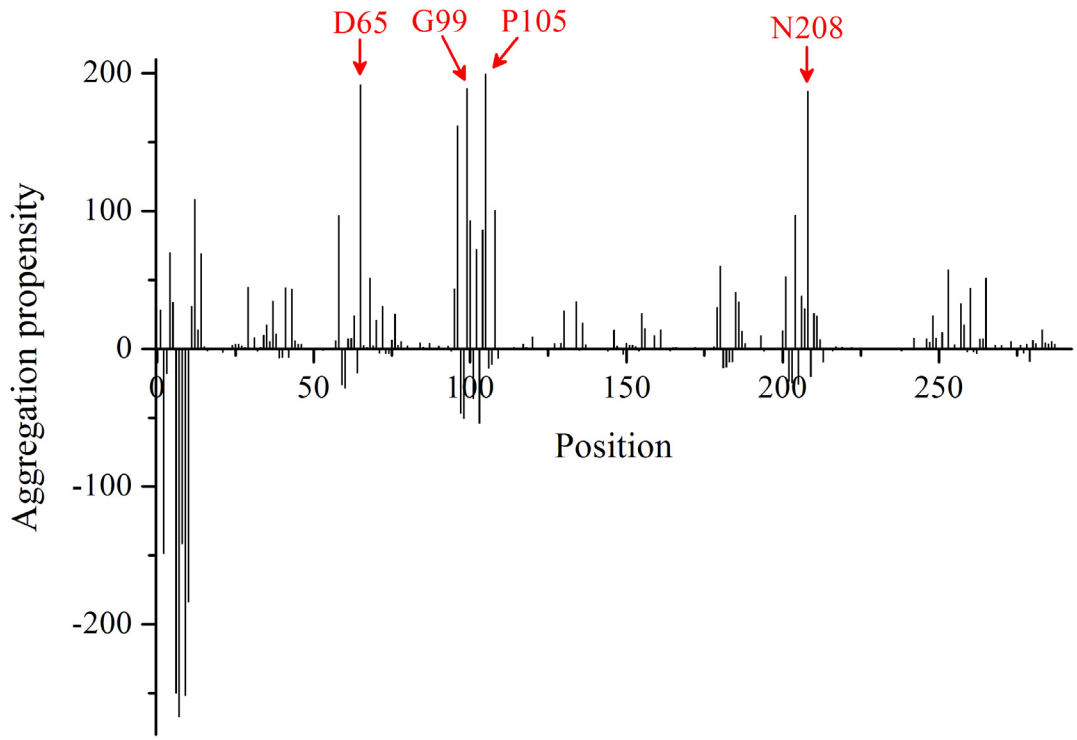


Figure 4.2 - The "aggregation propensity profile" of wild-type DHDPS according to the results of TANGO. The data indicate residues with the potential to, upon amino acid substitution, increase aggregation propensity of the protein as a whole.

The amino acids identified by TANGO as having an effect on the β-aggregation of DHDPS do not correlate to the result obtained from Zyggregator. Some of the general regions of interest are similar (*e.g.* positions 90-110 and positions 200-215) but the individual amino acids, identified by the two algorithms as having a distinct effect on the aggregation propensity of the protein as a whole, are dissimilar.

Although both TANGO and Zyggregator determine the aggregation propensity of the peptide sequences based on the properties of the amino acids (such as hydrophobicity, charge and influence on secondary structure), the results of the algorithms are different. The dissimilarities in the results likely arise from the differences in calculation as well as from the additional factors each consider. Hydrophilic/hydrophobic patterning is considered by Zyggregator and the thermodynamic properties of the protein structure and the intrinsic flexibility or existence of an unfolded state are considered by TANGO. Both algorithms demonstrate a good degree of accuracy upon comparison of the predictions with empirical observations of aggregation; however, they appear to be the most accurate for relatively short peptides [1, 2]. Empirical testing of Zyggregator was undertaken on a large globular protein (DHDPS), with the view of providing additional data to assist in the development of the algorithm.

4.3 RESULTS

4.3.1 PREDICTION OF THE AGGREGATION PROPENSITY OF *E. COLI* DHDPS BY ZYGGREGATOR

From the various algorithmic prediction programs available, the algorithm Zyggregator was chosen for experimental testing [2, 8]. The amino acids at positions 90, 207 and 270 were identified as having particular influence on the aggregation of *E. coli* DHDPS (see section 4.2). These residues are identified in the crystal structure of wild-type DHDPS shown in figure 4.3. From analysis of the positions in the crystal structure, it was apparent that the positions had few structural similarities. The glutamine at position 90 occurred in an α-helix on the exterior of the protein. The alanine at position 207 was also positioned in an α-helix but buried just below the surface of the protein. The proline at position 270 occurred in a loop very close to the lysine binding site, active site and the tight dimer interface (see chapter 1, section 1.10). Because of the potential involvement of this

residue in the catalytic site or the subunit interaction of the enzyme, an inter-species sequence comparison was carried out. This demonstrated that the amino acid residues at positions 90 and 207 display some variability; however, the proline at position 270 is highly conserved (figure 4.2). This suggests that the role of the proline at position 270 is critical to DHDPS from across a range of phyla.

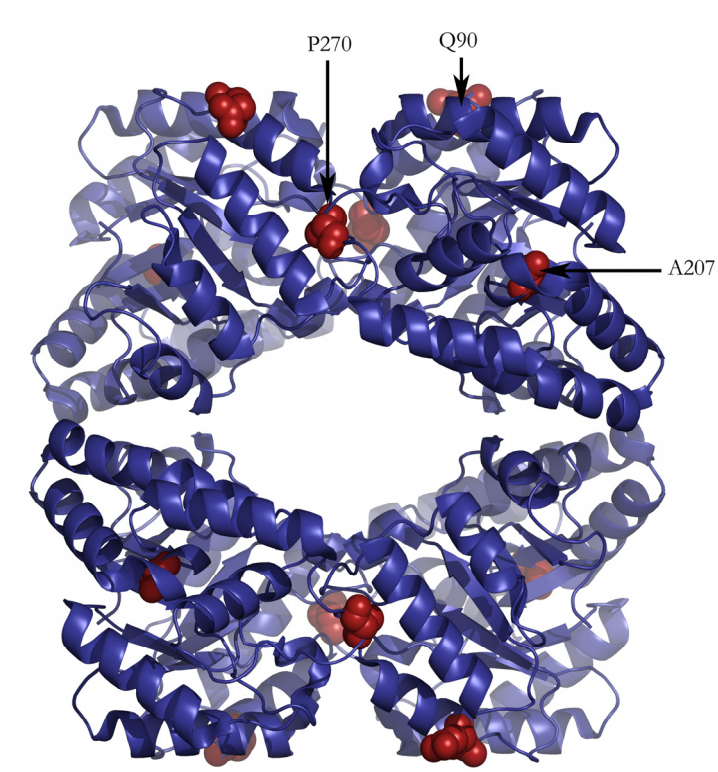


Figure 4.1 - The structure of wild-type *E. coli* DHDPS as published by Dobson *et al.*, 2005 [9]. The positions of Q90, A207 and P270 are shown in red. This image was created using PyMol [10] from the *E. coli* DHDPS PDB file (accession number 1yxc).

All three of the residues predicted by the algorithm to play an important role in aggregation propensity were substituted. The amino acids chosen to replace those at positions 90, 207 and 270 were identified by Zyggregator as having the greatest potential to increase the propensity of DHDPS to form β-sheet-specific aggregate (section 4.2). As a consequence of the predictions, glutamine at position 90 (Q90) was changed to a leucine (Q90L), alanine at position 207 (A207) was altered to a valine (A207V) and proline at position 270 (P270) was substituted with a threonine (P270T). Where the intrinsic aggregation potential was equal for several amino acids (figure 4.3), the replacement residues were chosen by identifying the amino acids whose codon was most similar to the wild-type residues, requiring a minimal change in the DNA sequence. The amino acid substitutions were

carried out using site directed mutagenesis using a commercial mutagenesis kit as described in chapter 8, section 8.2.14.

E. coli (Gram negative bacteria)	...L T Q R F N...V T A N V...R L P M T...
T. maritima (Gram positive bacteria)	...L V K Q A E...V V S N V...R L P L V...
M. tuberculosis (Gram positive bacteria)	...L A K A C A...V I A H L...R L P Q V...
S. elongatus (cyanobacteria)	...A T Q S A A...V A S H L...R L P L V...
N. tobacum (plant)	...A T E Q G F...V T S N L...R L P Y V...

Figure 4.2 - The amino acid sequences of DHDPS from various organisms. The red residue indicates the amino acid present at position 90 in the *E. coli* DHDPS sequence. The green residue represents position 207 and the blue residue represents position 270.

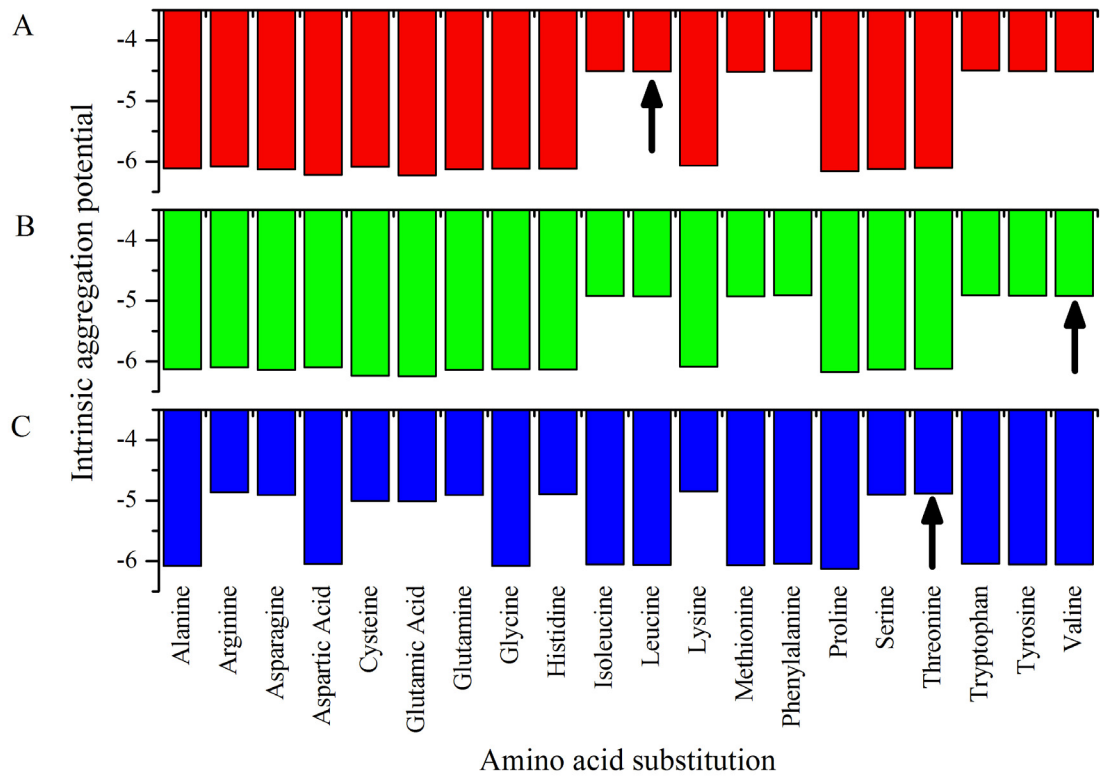


Figure 4.3 – The alteration of intrinsic aggregation propensity of DHDPS upon amino acid substitution for all the common amino acids: for **(A)** glutamine at position 90, **(B)** alanine at position 207 and **(C)** of the proline at position 270. As the intrinsic aggregation propensity becomes more positive, the aggregation propensity of the protein as a whole increases. The black arrows indicate the amino acids chosen to replace the wild-type residues.

4.3.2 GROWTH RATE AND EXPRESSION

Following confirmation of the mutagenesis through sequencing (chapter 8, section 8.2.14) the plasmids were transformed into the appropriate strains and the growth rate of each of the variants was assessed using an A_{600} assay. This simple analysis revealed interesting results. The growth rates of *E. coli* containing the plasmids with the mutated dapA genes (*i.e.* Q90L, A207V, and P270T) were significantly reduced compared to that containing the plasmid with the wild-type dapA gene (figure 4.4). This is most likely to be caused through the formation of inclusion bodies or other aggregate structures within the cells [11, 12]. It could be inferred from this, that P270T does increase the aggregation propensity of DHDPS. Unfortunately, the increased propensity to aggregate also results in a protein which is difficult to purify. Attempts to refold insoluble DHDPS P270T using high concentrations of urea were unsuccessful. This was attributed to the protein's size and is congruent with the results of the CD spectroscopy (described in chapter 2, section 2.3.2.6 and appendix 1, section A1.2) which showed that DHDPS did not regain significant secondary structure upon thermal denaturation and attempted renaturation, suggesting that the refolding of DHDPS is non-trivial *in vitro*.

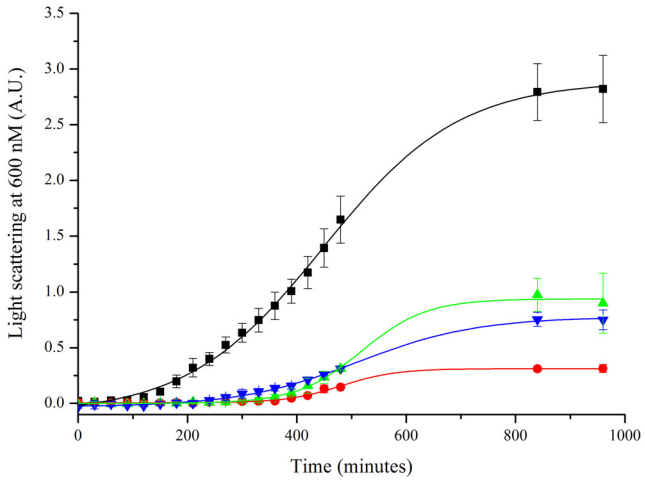


Figure 4.4 - The growth rates of *E. coli* XL1 Blue containing the dapA gene on the pJG001 plasmid (black), pJG001 Q90L (green), pJG001 A207V (blue) and pJG001 P270T (red). The error bars represent the SEM of three replicates. A sigmoid (Boltzmann) distribution was used for the line of best fit.

Multiple attempts to purify DHDPS P270T failed (see SDS PAGE results in figure 4.5), thus no further characterisation was performed on this variant. The impact of the A207V and Q90L amino

acid substitutions were not as detrimental, thus the proteins were purified and characterised. Due to the significant effects of the polyhistidine tags on the amorphous and β-sheet-specific aggregation propensities of wild-type DHDPS (see chapter 3), the combinatorial effect of the polyhistidine tags with the proteins containing the amino acid substitutions Q90L and A207V were also tested. The combinatorial effect of the amino acids and the predicted amino acid substitutions were assessed using Zyggregator [2, 8]. The results indicated that all the DHDPS Q90L and A207V variants exhibit increased aggregation potentials (quantified by the Z_{agg} score). These increased aggregation scores were attributed to the increase in the number of hydrophobic/hydrophilic patterns arising from the presence of the amino acid substitutions at positions 90 and 207. The combinatorial affects of the polyhistidine tags reduced the predicted aggregation propensity. Comprehensive results of the Zyggregator analyses are presented in appendix 3, section A3.1.

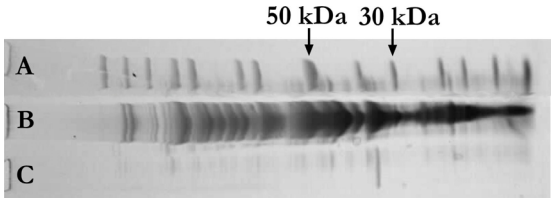


Figure 4.5 – SDS-PAGE. The lanes were loaded with **(A)** molecular weight marker with relevant bands indicated, **(B)** crude sample containing DHDPS P270T, **(C)** purified P270T (post ion exchange and hydrophobic interaction chromatography as described in chapter 8, section 8.4.5 and 8.4.6).

4.3.3 PURIFICATION

The purification of *E. coli* DHDPS Q90L and A207V was carried out according to the method adapted from Dobson *et al.* 2004 [13] as detailed in chapter 8, sections 8.4.1, 8.4.4 – 8.4.8. Following purification by ion exchange chromatography (IEX) and hydrophobic interaction chromatography (HIC) the average yield for DHDPS Q90L was 28.0 mg (79 units with a specific activity of 2.83 units/mg) and 17.7 mg for DHDPS A207V (46 units with a specific activity of 2.63 units/mg) per litre of culture. The yield for wild-type DHDPS averaged was 19.6 mg (76 units with a specific activity of 3.83 units/mg) per litre of culture. The amino acid substitution at position 90 increased the yield compared to wild-type; however, the specific activity was reduced. The yield for DHDPS A207V was reduced, also exhibiting a lower specific activity.

The addition of the polyhistidine tags to DHDPS Q90L resulted in increased yield; pET M11 DHDPS Q90L yielded an average of 42.1 mg (121.9 units with a specific activity of 2.90 units/mg) per litre of culture and pET 151/D-TOPO DHDPS Q90L yielded an average 39.2 mg (126 units with a specific activity of 3.26 units/mg) per litre of culture. Cleavage of the pET 151/D-TOPO polyhistidine tag from the DHDPS Q90L was undertaken as per the protocol detailed in chapter 8, section 8.4.9. Some protein was lost during this step due to the incomplete cleavage of all the polyhistidine tags from the protein and the subsequent affinity purification. The cleavage reaction of the pET 151/D-TOPO DHDPS Q90L yielded an average 66% of the original protein present. The yield following cleavage averaged 25.7 mg (68 units with a specific activity of 2.66 units/mg) per litre of culture, a slight reduction in yield and specific activity compared to the yield obtained from the non-polyhistidine-tagged species. Qualitative analysis of the SDS PAGE gel (figure 4.6) shows that cleaved pET 151/D-TOPO DHDPS Q90L contains fewer contaminants than DHDPS Q90L. This suggests that the reduction in specific activity is due to a change in the protein's biophysical parameters and not a reduced level of purity of cleaved pET 151/D-TOPO DHDPS Q90L consistent with the impact of this tag on the wild-type enzyme (chapter 3, section 3.3.3.1). As well as facilitating the rapid purification of DHDPS Q90L, the addition of the polyhistidine tags improved the purity of most of the protein preparations.

The addition of the polyhistidine tags to DHDPS A207V had mixed effects. Following the addition of the pETM11 polyhistidine tag to DHDPS A207V, the yield was reduced compared to DHDPS A207V. The average yield was 12.3 mg (27 units with a specific activity of 2.21 units/mg) per litre of culture. The yield upon addition of the pET 151/D-TOPO polyhistidine tag was marginally improved, yielding an average 14.1 mg (36 units with a specific activity of 2.53 units/mg) per litre of culture. The cleavage reaction of the pET 151/D-TOPO polyhistidine tag resulted in a further loss of protein, yielding a mean 8.3 mg (21 units with a specific activity of 2.54 units/mg) per litre of culture. The SDS PAGE gel analysis revealed that the purity of all the polyhistidine-tagged variants is higher than DHDPS A207V (figure 4.6). Because the addition of the polyhistidine tags to DHDPS A207V significantly reduced the yields and specific activities, it was hypothesised that the presence of the polyhistidine tag in combination with the amino acid change at position 207 affected the biophysical properties of the protein as a whole. This may have been due to reduced stability or an increased propensity to aggregate. The results of the experiments designed to test this hypothesis are described below (section 4.3.4).

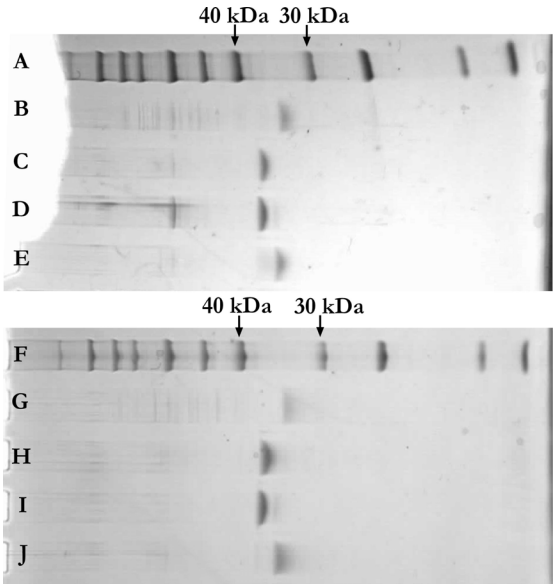


Figure 4.6 – SDS-PAGE. The lanes were loaded with **(A)** molecular weight marker with relevant bands indicated, **(B)** DHDPS Q90L, **(C)** pET M11 DHDPS Q90L, **(D)** pET 151/D-TOPO DHDPS Q90L, **(E)** cleaved pET 151/D-TOPO DHDPS Q90L, **(F)** molecular weight marker with relevant bands indicated, **(G)** DHDPS A207V, **(H)** pET M11 DHDPS A207V, **(I)** pET 151/D-TOPO DHDPS A207V and **(J)** cleaved pET 151/D-TOPO DHDPS A207V.

As was the case with the wild-type DHDPS and its polyhistidine-tagged variants, the protein samples were stored at -20 ºC and used within a 6 month period of purification [14, 15]. Where possible, all assays were carried out on protein obtained from a single purification.

4.3.4 BIOPHYSICAL PROPERTIES

The biophysical techniques used to establish the properties of the DHDPS point mutants are described in chapters 3 and 4. All the point mutants were compared to wild-type DHDPS.

4.3.4.1 KINETICS

DHDPS Q90L and pET 151/D-TOPO DHDPS Q90L exhibited similar k_{cat}^{app} values to those of the wild-type enzyme (table 4.3). The pET M11 polyhistidine tag caused a slight decrease in k_{cat}^{app} as did the residues remaining following the cleavage of the polyhistidine tag from pET 151/D-TOPO DHDPS Q90L, consistent with the results obtained for the pET 151/D-TOPO tagged wild-type

variants (see chapter 3, section 3.3.3.1). The amino acid substitution at position 90 appeared to cause an increase in the K_m^{app} with respect to pyruvate and (*S*)-ASA.

DHDPS A207V and its polyhistidine-tagged variants generally had k_{cat}^{app} parameters lower than wild-type and the pET M11 DHDPS A207V exhibited a k_{cat}^{app} outside the published range of k_{cat} values [13, 16, 17]. The DHDPS A207V variants exhibited elevated K_m for both pyruvate and (*S*)-ASA.

	k_{cat}^{app} with respect to Pyruvate (s^{-1}) ± SEM	**$K_{m\ Pyr}^{app}$ (mM) ± SEM**	**k_{cat}^{app} with respect to (*S*)-ASA (s^{-1}) ± SEM**	**$K_{m\ (S)\text{-}ASA}^{app}$ (mM) ± SEM**
Wild-type DHDPS	118 ± 3	0.32 ± 0.03	123 ± 5	0.14 ± 0.02
DHDPS Q90L	94 ± 4	0.60 ± 0.10	83 ± 4	0.27 ± 0.05
pET M11 DHDPS Q90L	80 ± 5	0.53 ± 0.11	101 ± 5	0.24 ± 0.04
pET 151/D-TOPO DHDPS Q90L	101 ± 5	0.54 ± 0.10	102 ± 6	0.44 ± 0.07
Cleaved pET 151/D-TOPO DHDPS Q90L	59 ± 3	0.66 ± 0.13	107 ± 6	0.31 ± 0.05
DHDPS A207V	73 ± 2	0.25 ± 0.03	91 ± 4	0.17 ± 0.02
pET M11 DHDPS A207V	82 ± 2	0.37 ± 0.04	56 ± 1	0.19 ± 0.02
pET 151/D-TOPO DHDPS A207V	96 ± 5	0.47 ± 0.09	62 ± 2	0.25 ± 0.03
Cleaved pET 151/D-TOPO DHDPS A207V	95 ± 4	0.33 ± 0.06	64 ± 3	0.22 ± 0.03

Table 4.3 – The kinetic parameters for the variants of DHDPS compared to those of wild-type. The k_{cat} values for all the polyhistidine-tagged variants are generally within the range of published values (k_{cat} ranging from 79 to 188 ± 7 **s^{-1}**). The apparent Michaelis constant for pyruvate is higher than the published values (K_{mPyr} ranging from 0.1 ± 0.01 mM to 0.26 ± 0.03 mM). The apparent Michaelis constant for ASA of all polyhistidine-tagged variants were higher than that of the wild-type ($K_{m\ (S)\text{-}ASA}$ ranging from 0.11 ± 0.01 mM to 0.13 ± 0.02 mM) [13, 16, 17]. The two k_{cat} values arise from the use of pseudo-single substrate models rather than a complete two substrate analysis. The data presented are the means of at least three replicates ± SEM.

Both the amino acid substitutions at position 90 and 207 appeared to alter the kinetic parameters of DHDPS. As the ability of DHDPS to catalyse the condensation of (*S*)-ASA and pyruvate to form DHDP was dependent on an unchanged active site, the variability observed in *E. coli* DHDPS Q90L and A207V and their polyhistidine-tagged variants may have been a result of an aberrant tertiary or quaternary structure. These possibilities were investigated and are detailed in the following sections.

4.3.4.2 CRYSTALLOGRAPHIC ANALYSIS

Attempts were made to ascertain the effect of both the mutations made at positions 90 and 207 on the structure of DHDPS through crystallographic analysis. Un-tagged DHDPS A207V was

crystallised using the technique described in Mirwaldt *et al.* 1995 [18] and as modified by Dobson *et al.* 2005 [9]. The crystals were grown at 12 ºC and appeared after approximately 3 days and were similar in size and shape to those of wild-type DHDPS [9]. The refinement statistics are shown in table 4.4. Repeated attempts were made to crystallise un-tagged DHDPS Q90L under the same conditions and at 4 ºC; however, these attempts were unsuccessful. In addition to the standard conditions, sparse-matrix screens (Hampton Research screens HR2-110 and HR2-112) were employed in an attempt to find more favorable conditions. These yielded no promising results. Attempts were also made to crystallise cleaved pET 151/D-TOPO DHDPS Q90L using the conditions under which wild-type DHDPS forms crystals. These conditions proved unsuccessful as did subsequent sparse matrix screens. The inability to obtain DHDPS Q90L crystals is an interesting result because it suggests that the equilibrium involved in crystal formation (chapter 1, figure 1.2) has been perturbed for this variant.

Resolution (data processing) (Å)		2.2
Number of images		233
Ocillation range (°)		0.5
Space group [a]		$P3_121$
Unit-cell parameters (Å)		a = 121.21
		b = 121.21
		c = 110.19
Number of reflections/unique reflections		203685/45593
Completeness		99.3 (99.3)
R_{merge} [b]		0.119 (0.275)
$I/\sigma(I)$		4.6
Resolution (refinement) (Å)		2.28 - 2.20
R_{free} [c]		0.235
R_{cryst} [d]		0.189
Mean B value (Å²)		22.8
Estimated coordinate error		0.182
Residues/solvent molecules		584/585
rmsd from ideal geometry	Bond lengths (Å)	0.008
	Bond angles (°)	1.28

[a] Wild-type space group was P3121 and cell dimensions were a=b=122.41 Å, c=111.22 Å
[b] R_{merge} = SUM (ABS(I - [I]))/SUM(I)
[c] R_{free} is based on 5 % of the total reflections
[d] $= \sum \left| \left| F_{obs} \right| - \left| F_{calc} \right| \right| / \sum \left| F_{obs} \right|$

Table 4.4 – Data set and refinement statistics. The values in parentheses are for the highest resolution shell.

The crystals of DHDPS A207V diffracted to 2.2 Å and showed no significant changes to the secondary, tertiary or quaternary structures when compared to wild-type DHDPS (figures 4.7 – 4.9 are superpositions of the wild-type protein with the structure obtained for DHDPS A207V). This led to the hypothesis that any changes in the kinetic properties can be attributed to the effect of the amino acid substitution on the protein in a non-crystalline state. As crystallographic data only provides information regarding the protein in its crystal state, the behavior of both DHDPS Q90L and A207V in solution, were investigated and are described in the following sections.

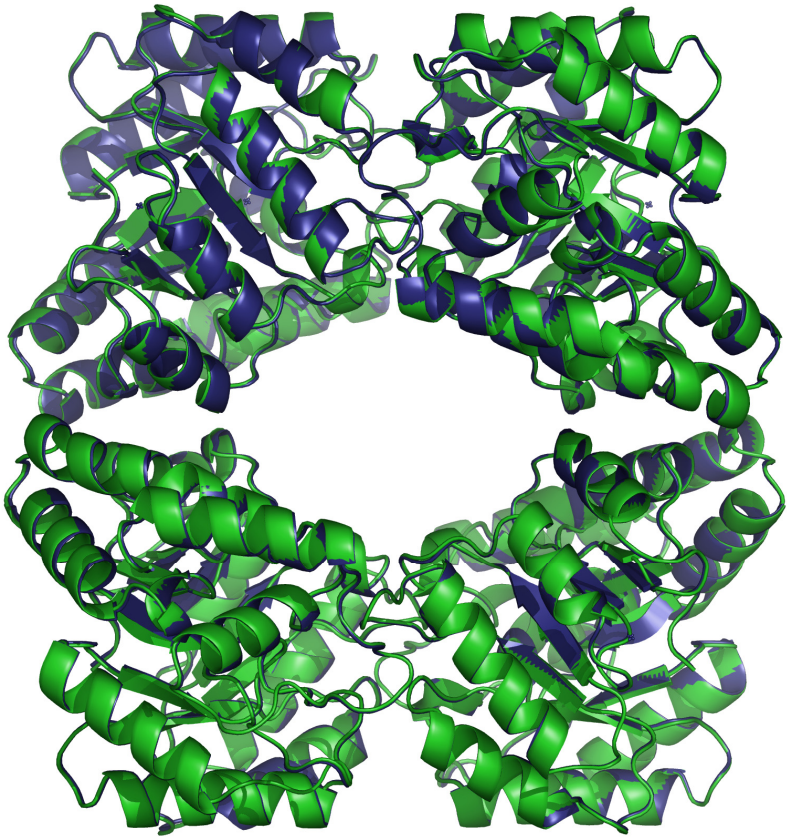


Figure 4.7 – Alignment of the structure of wild-type DHDPS (blue) with the structure of DHDPS A207V (green) demonstrating no significant alterations to either the secondary, tertiary or quaternary structures. The rmsd for the tetramer = 0.268 Å. This image was created using PyMol [10] from the *E. coli* DHDPS PDB file (accession number 1yxc) [9] and the crystal structure of DHDPS A207V.

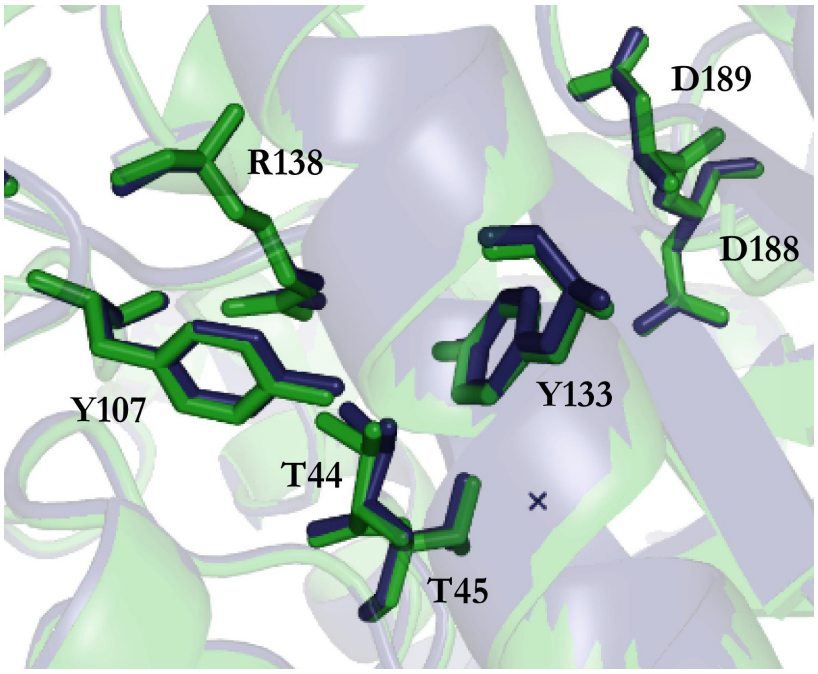


Figure 4.8 – Alignment of the structure of wild-type DHDPS (blue) with the structure of DHDPS A207V (green). Superposition of the residues at the active site indicates no change in the surrounding structure. This image was created using PyMol [10] from the *E. coli* DHDPS PDB file (accession number 1yxc) [9] and the crystal structure of DHDPS A207V.

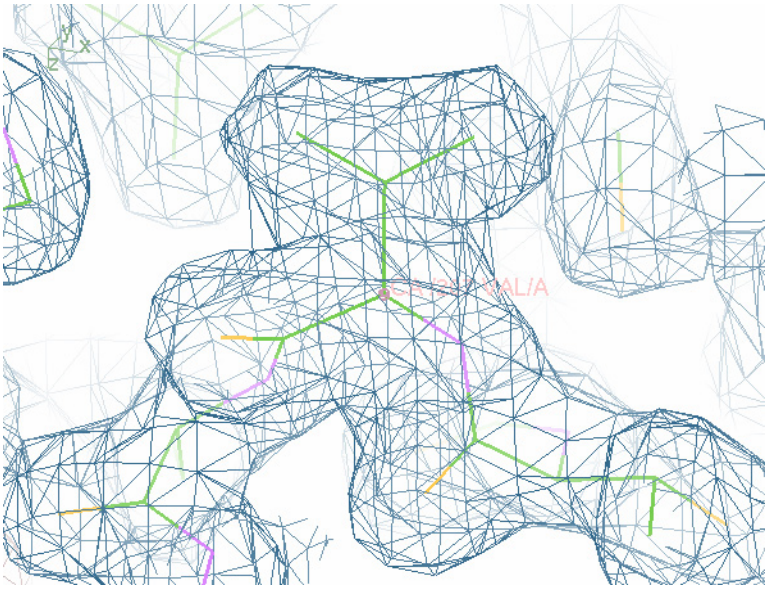


Figure 4.9 – Residue 207 of DHDPS A207V showing electron density. The 2Fo-Fc electron density (blue) is contoured to 1 σ. This image was created using COOT [19].

4.3.4.3 ANALYTICAL GEL PERMEATION LIQUID CHROMATOGRAPHY

Analytical gel permeation liquid chromatography was used to assess the quaternary structures of DHDPS Q90L, A207V and their polyhistidine variants in solution. The results were compared to that of wild-type DHDPS. Figure 4.10 and 4.11 show the chromatograms obtained for each of the DHDPS variants.

The molecular masses of DHDPS Q90L, pET M11 DHDPS Q90L, pET 151/D-TOPO DHDPS Q90L, cleaved pET 151/D-TOPO DHDPS Q90L, DHDPS A207V, pET 151/D-TOPO DHDPS A207V, and cleaved pET 151/D-TOPO DHDPS A207V calculated from the elution volumes were generally within the 10 % error of the expected molecular masses [20, 21] (see table 4.5). pET M11 DHDPS A207V displayed a slightly more aberrant calculated molecular mass (109 kDa) which suggested the protein could exist as either a trimer or a tetramer. Because of this, further investigation of the quaternary structure was carried out by AUC and is described in section 4.3.4.4. The observed inconsistencies may be a result of the proteins behaving in a non-ideal manner whilst on the column; however, most results still support the assertion that all the variants exist as tetramers. These results indicate that attaching a polyhistidine tag had no significant effect on the quaternary structure of DHDPS Q90L or A207V, corroborating the results for the tagged wild-type enzymes (chapter 3, section 3.3.3.3). Although the proteins were all 0.5 mg/mL, variability in the mAU was observed. This may be due to errors in calculation of protein concentration or a result of the protein aggregating on the column. This latter hypothesis is investigated further in section 4.3.4.7.

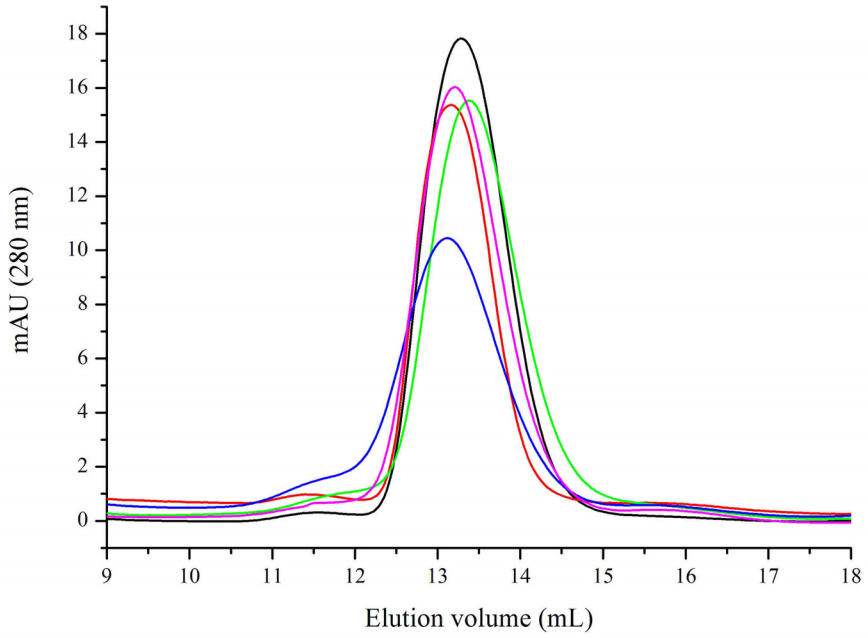


Figure 4.10 – Elution profiles for wild-type DHDPS (black), DHDPS Q90L (red), pET M11 DHDPS Q90L (green), pET151/D-TOPO DHDPS Q90L (blue), and cleaved pET151/D-TOPO DHDPS Q90L (pink). The chromatograms were produced by measuring absorbance at 280 nm as a function of elution volume (V_e).

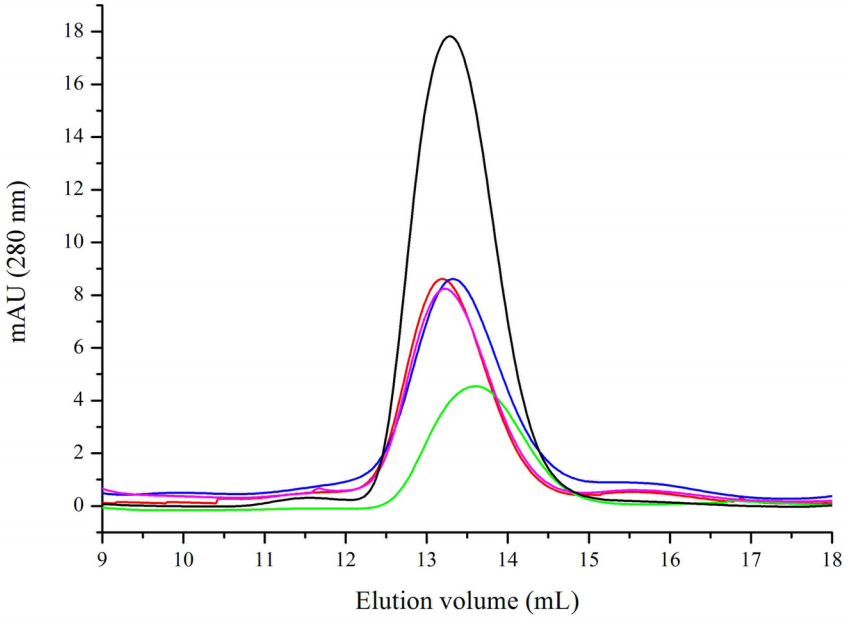


Figure 4.11 – Elution profiles for wild-type DHDPS (black), DHDPS A207V (red), pET M11 DHDPS A207V (green), pET151/D-TOPO DHDPS A207V (blue) and cleaved pET151/D-TOPO DHDPS A207V (pink). Data indicates that all the variants exist as tetramers. The chromatograms were produced by measuring absorbance at 280 nm as a function of elution volume (V_e).

Variant	Predicted molecular mass	Calculated molecular mass
Wild-type DHDPS	125 kDa	134 kDa
DHDPS Q90L	125 kDa	143 kDa
pET M11 DHDPS Q90L	138 kDa	126 kDa
pET151/D-TOPO DHDPS Q90L	140 kDa	139 kDa
cleaved pET151/D-TOPO DHDPS Q90L	128 kDa	146 kDa
DHDPS A207V	125 kDa	141 kDa
pET M11 DHDPS A207V	138 kDa	109 kDa
pET151/D-TOPO DHDPS A207V	140 kDa	131 kDa
cleaved pET151/D-TOPO DHDPS A207V	128 kDa	139 kDa

Table 4.5 – The molecular masses of wild-type DHDPS and its variants as predicted by "Protein Calculator" (http://www.scripps.edu/~cdputnam/protcalc.html) and as calculated from the calibration curve shown in chapter 2, section 2.3.2.3.

4.3.4.4 ANALYTICAL ULTRA-CENTRIFUGATION

Sedimentation velocity AUC analysis of DHDPS was performed on pET M11 DHDPS Q90L and A207V in order to confirm the results of the analytical gel permeation liquid chromatography and provide further information on the thermodynamic and hydrodynamic properties of the protein (section 8.6.2). The sedimenting boundaries for both variants did not undergo any significant spreading with respect to time thus suggesting the samples were largely homogeneous (figures 4.12(B) and 4.13(B)). Confirmation of this was achieved by fitting the data to continuous size-distribution models [22, 23]. The $c(s)$ data showed primary peaks with an ordinate maxima corresponding to a sedimentation coefficient of 6.7 S and 6.6 S for pET M11 Q90L and pET M11 A207V respectively (figures 4.12(C) and 4.13(C)). These values were congruent with those obtained for wild-type DHDPS (chapter 2, section 2.3.2.4) and with values from existing literature for tetrameric DHDPS [17, 24, 25]. The quality of the nonlinear least squares best-fit was affirmed by the low rmsd numbers of 0.0045 and 0.0042 and the random residual distribution data (figures 4.12(A) and 4.13(A)). The small secondary peak in the A207V sample is similar to the results obtained for wild-type enzyme in this study. This may be a contaminant; however as SDS-PAGE shows a relatively high level of purity (see figure 4.6), it is likely that this peak represents an equilibrium of dimers and tetramers as was described in Perugini *et al.* [25] as discussed in chapter 2, section 2.3.2.4. The presence of the secondary peak for wild-type is not consistent in published literature [17, 25, 26]. Therefore, it is unlikely that the second peak observed for A207V reflects a difference in stability compared to Q90L.

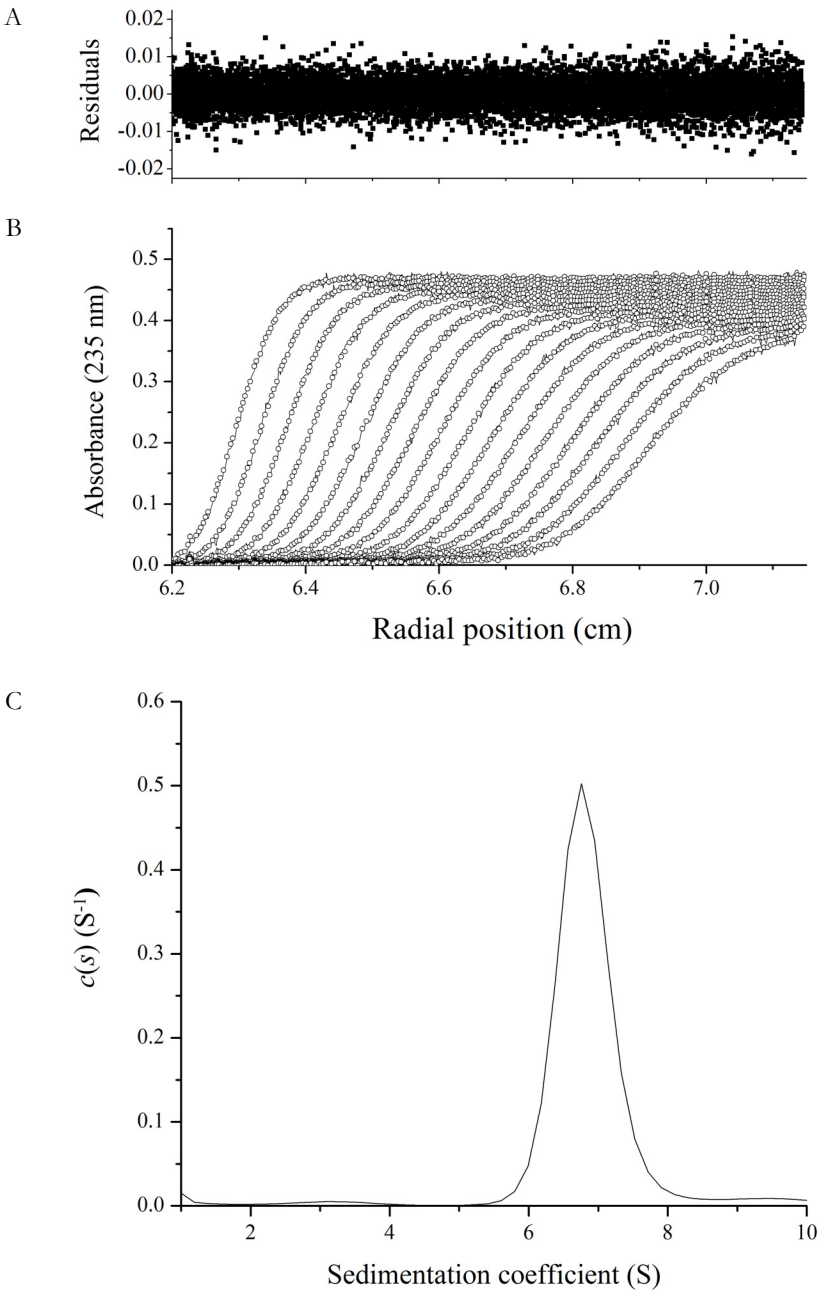


Figure 4.12 – The sedimentation velocity of pET M11 DHDPS Q90L carried out at 0.1 mg/mL. **(A)** Residuals for the nonlinear least squares best fit are plotted as a function of the radial position from the axis of rotation (cm). **(B)** Absorbance at 235 nm plotted as a function of the radial position from the axis of rotation (cm). Both the raw data (symbols) and fitted data (solid line) are plotted at time intervals of 6 minutes. Raw data are overlaid with the nonlinear least squares best fit to a continuous size-distribution model. **(C)** The continuous sedimentation coefficient ($c(s)$) distribution plotted as a function of the sedimentation coefficient (S). The fit obtained was using a resolution of 100 species between s_{min} of 1 S and s_{max} of 10 S with P = 0.95, $\bar{\nu}$ = 0.7378, ρ = 1.005 g/mL, η = 1.0214 cp and f/f_0 = 1.62.

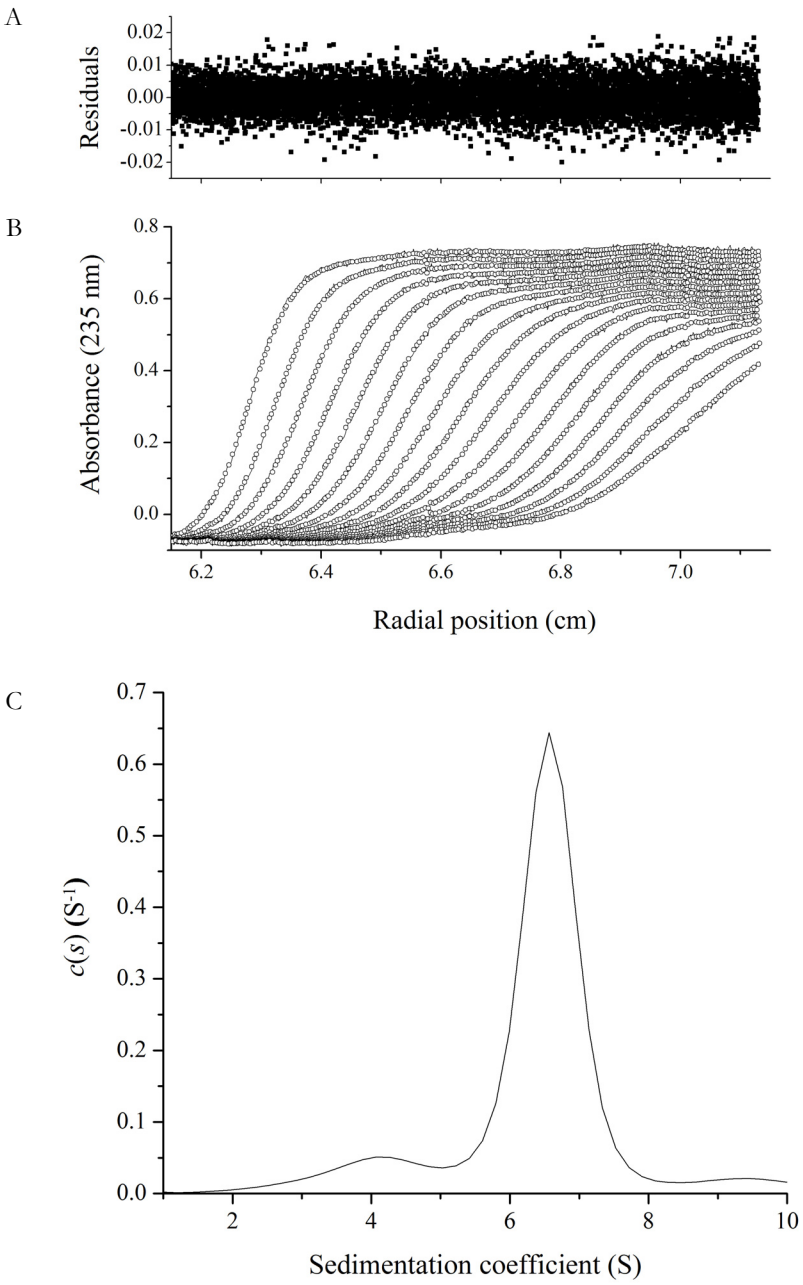


Figure 4.13 – The sedimentation velocity of pET M11 DHDPS A207V carried out at 0.1 mg/mL. **(A)** Residuals for the nonlinear least squares best fit are plotted as a function of the radial position from the axis of rotation (cm). **(B)** Absorbance at 235 nm plotted as a function of the radial position from the axis of rotation (cm). Both the raw data (symbols) and fitted data (solid line) are plotted at time intervals of 6 minutes. Raw data are overlaid with the nonlinear least squares best fit to a continuous size-distribution model. **(C)** The continuous sedimentation coefficient ($c(s)$) distribution plotted as a function of the sedimentation coefficient (S). The fit obtained was using a resolution of 100 species between s_{min} of 1 S and s_{max} of 10 S with P = 0.95, $\bar{\nu}$ = 0.7382, ρ = 1.005 g/mL, η = 1.0214 cp and f/f_0 = 1.55.

The results of the AUC confirmed the results obtained by analytical gel permeation chromatography. In solution both DHDPS Q90L and DHDPS A207V exist as tetramers. As the analytical gel permeation liquid chromatography showed no difference between the non-polyhistidine-tagged variants and their polyhistidine-tagged counterparts, and due to the time consuming nature of the AUC assays, the polyhistidine-tagged variants of DHDPS Q90L and A207V were not assessed using this technique.

4.3.4.5 DIFFERENTIAL SCANNING FLUORIMETRY

The thermal denaturation temperatures of DHDPS Q90L, DHDPS A207V and their polyhistidine-tagged variants were assessed using a thermofluor technique as described in chapter 2, section 2.1.5. The results were compared to those obtained for wild-type DHDPS (chapter 2, section 2.3.2.5). Statistical analysis was carried out as described in chapter 3, section 3.3.3.4. The analysis of deviance table for all DHDPS variants is presented in appendix 1, table A1.2.

The effect of salt and pH on the thermal denaturation of the DHDPS Q90L in 100 mM phosphate buffer is shown in figure 4.14 and summarised in figure 4.15. The data for Q90L exhibited high initial fluorescence characteristic of DHDPS at pH 2. Despite repeated attempts, using fresh protein from several purifications, the data always exhibited this trend. This pattern of fluorescence was observed under all conditions thus was unlikely to be an effect of the buffer and was thought to be a result of the protein existing in a partially unfolded state prior to the assay [27]. This will be discussed in section 4.3.4.6.

DHDPS Q90L variants underwent normal thermal denaturation, with low initial fluorescence, increasing with elevation of the temperature (see appendix A3, figure A3.1). Upon comparison of the polyhistidine-tagged variants of DHDPS Q90L, the thermal denaturation temperature was not significantly different except for cleaved pET 151/D-TOPO DHDPS Q90L which, at pH 11, exhibited a significantly lower thermal denaturation temperature compared to the polyhistidine-tagged variants of DHDPS Q90L. Comparison of the DHDPS Q90L variants with wild-type DHDPS demonstrated no significant differences except for at pH 4.5. At this pH, the polyhistidine-tagged variants underwent thermal denaturation at a higher temperature than wild-type. This may

indicate that these proteins are more stable than wild-type. The inconsistent values obtained for the maximum RFU are discussed in chapter 2, section 2.3.2.5.

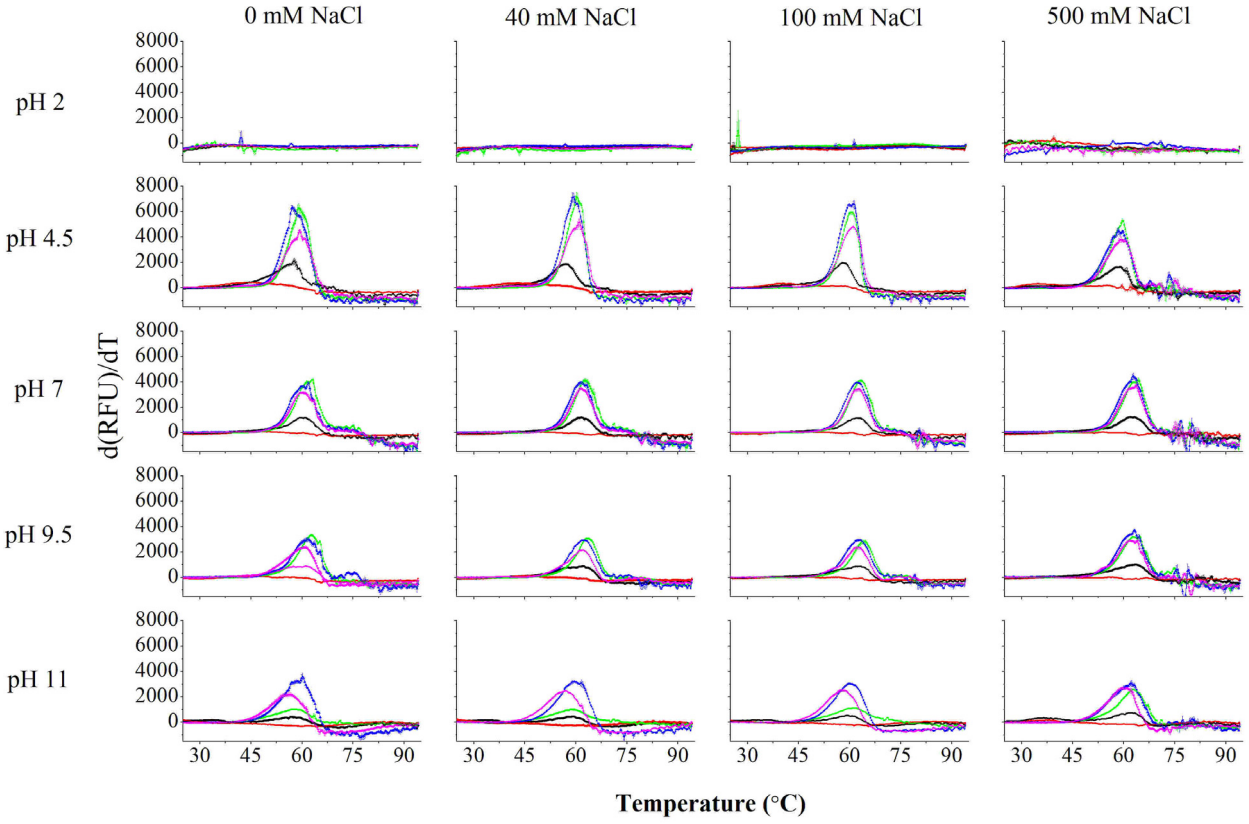


Figure 4.14 – The thermal denaturation profiles of wild-type DHDPS (black), DHDPS Q90L (red), pET M11 DHDPS Q90L (green), pET 151/D-TOPO DHDPS Q90L (blue) and cleaved pET 151/D-TOPO DHDPS Q90L (pink) in 100 mM phosphate buffer (pH 7, 40 mM NaCl) as monitored by SYPRO Orange fluorescence. The data plotted are the derivatives of the increase in fluorescence monitored by the BioRad IQ5 (presented in appendix A3, figure A3.1). The peaks indicate the temperature at which the protein is unfolding the most rapidly and is recorded as the mean thermal denaturation temperature. The data plotted are the mean values of three replicates ± SEM.

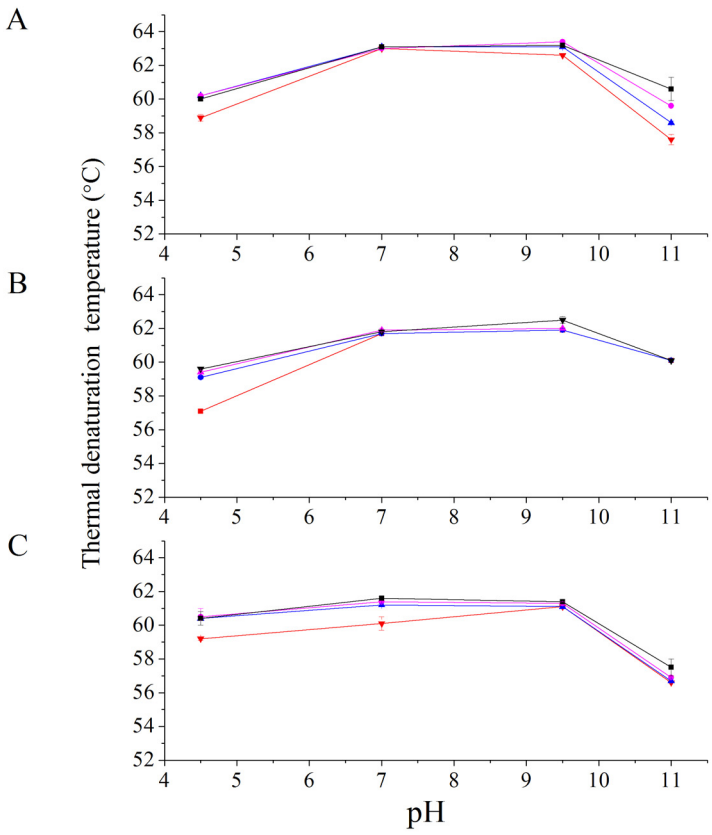


Figure 4.15 – The primary thermal denaturation temperatures of **(A)** pET M11 DHDPS Q90L, **(B)** pET 151/D-TOPO DHDPS Q90L and **(C)** cleaved pET 151/D-TOPO DHDPS Q90L in 100 mM phosphate buffer at pH 4.5, pH 7, pH 9.5 and pH 11 with 0 mM (green), 40 mM (blue), 100 mM (pink) and 500 mM NaCl (black) as monitored by SYPRO Orange fluorescence and as assessed from the derivation of the fluorescence data. These results are summarised in appendix 3, section A3.2.

Some variants demonstrated biphasic data under some conditions (*e.g.* pH 9.5, 0 mM NaCl); however, the following discussion is concerned with the primary thermal denaturation temperature peak. The biphasic data will be discussed later in the chapter. As described for wild-type DHDPS and its tagged variants (chapter 2, section 2.3.2.5 and chapter 3, section 3.3.3.4), the effect of salt on the thermal denaturation temperatures of the polyhistidine-tagged variants was not significant. At pH 2, all the DHDPS Q90L variants exhibited high initial fluorescence suggesting that hydrophobic patches were exposed prior to the addition of the dye [27]. As CD spectra suggested that wild-type DHDPS retained some secondary structure at pH 2 (see section 2.3.2.6 and appendix A1, section A1.3), the high initial fluorescence was unlikely to be due to complete denaturation of the protein. As observed for the wild-type protein, the DHDPS Q90L variants were most thermally stable at pH 7 and pH 9.5. The pI values for all the proteins were similar to wild-type and its polyhistidine-tagged

variants (6.44 for wild-type DHDPS, 6.44 for DHDPS Q90L, 6.57 for pET M11 DHDPS Q90L, 6.57 for pET 151/D-TOPO DHDPS Q90L and 6.27 for cleaved pET 151/D-TOPO DHDPS Q90L). Thus the high thermal stability at pH 7 and pH 9.5 is most likely to be due to the deprotonation of the proteins causing an increase in solubility and thermal stability [28, 29]. At pH 11 and 4.5 the thermal denaturation temperatures of the DHDPS Q90L variants were reduced, as observed for the wild-type enzyme.

The effect of salt and pH on the thermal denaturation of the DHDPS A207V in 100 mM phosphate buffer is shown in figure 4.16 and figure 4.17. Comparison of the polyhistidine-tagged variants of DHDPS A207V revealed that thermal denaturation temperatures were significantly affected by the presence of the amino acid substitution and the various polyhistidine tags. This effect is particularly notable at pH 4.5 (see figure 4.16 and 4.17). At this pH, all the polyhistidine-tagged variants of DHDPS A207V exhibit significantly higher thermal denaturation temperatures than both wild-type and DHDPS A207V. At higher pH, the pET 151/D-TOPO polyhistidine-tagged variant has a lower thermal denaturation temperature than wild-type and the other A207V variants. The results suggest that the sequence of the polyhistidine motif and pH have an impact on the thermal denaturation of DHDPS containing the amino acid substitution at position 207.

As described for DHDPS Q90L, the effect of salt on the thermal denaturation temperatures of DHDPS A207V was less significant than the effect of pH. At pH 2, all proteins containing the amino acid substitution at positions 207, like other DHDPS variants, exhibited high initial fluorescence suggesting that hydrophobic patches were exposed prior to the addition of the dye [27]. The DHDPS A207V variants were most thermally stable at pH 7 and pH 9.5 except for DHDPS A207V which exhibited a lower thermal denaturation temperature at pH 9.5. The pI values for all the proteins were similar to wild-type and its polyhistidine-tagged variants (6.44 for wild-type DHDPS, 6.44 for DHDPS A207V, 6.57 for pET M11 DHDPS A207V, 6.57 for pET 151/D-TOPO DHDPS A207V and 6.27 for cleaved pET 151/D-TOPO DHDPS A207V). At pH 7 and pH 9.5, deprotonation of the proteins probably results in an increase in solubility and therefore thermal stability [28, 29]. At pH 11 and 4.5 the thermal denaturation temperature of the DHDPS A207V variants were reduced, as observed for all other DHDPS variants.

Some DHDPS Q90L and A207V polyhistidine-tagged variants appeared to undergo two thermal denaturation events under certain conditions (see figure 4.18 and appendix A3, tables A3.4 and

A3.5). Previous reports of biphasic data have attributed the presence of two thermal denaturation temperatures to the presence of two rate-determining unfolding processes [30]. In the case of DHDPS, this could be attributed to an abnormal protein fold or the presence of a protein species with an aberrant quaternary structure. The former can be investigated using circular dichroism (CD) spectroscopy (see section 4.3.4.6). However, as these second peaks exhibited higher thermal denaturation temperatures than that of the wild-type protein under comparable conditions it is likely that they represent a species with an aberrant quaternary structure or a population of aggregate. This will be further investigated in section 4.3.4.7.

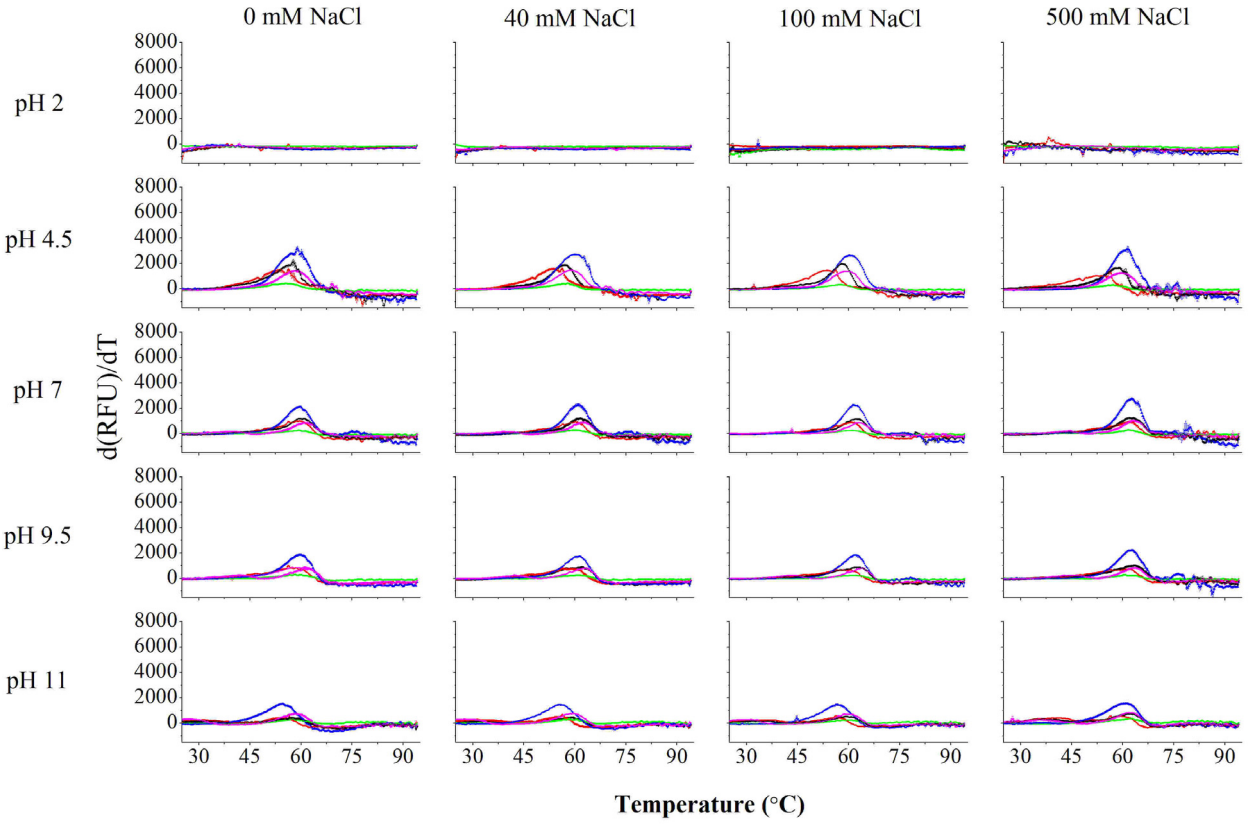


Figure 4.16 – The thermal denaturation profiles of wild-type DHDPS (black), DHDPS A207V (red), pET M11 DHDPS A207V (green), pET 151/D-TOPO DHDPS A207V (blue) and cleaved pET 151/D-TOPO DHDPS A207V (pink) in 100 mM phosphate buffer (pH 7, 40 mM NaCl) as monitored by SYPRO Orange fluorescence. The data plotted are the derivatives of the increase in fluorescence monitored by the BioRad IQ5 (presented in appendix A3, figure A3.2). The peaks indicate the temperature at which the protein is unfolding the most rapidly and is recorded as the mean thermal denaturation temperature. The data plotted are the mean values of three replicates ± SEM.

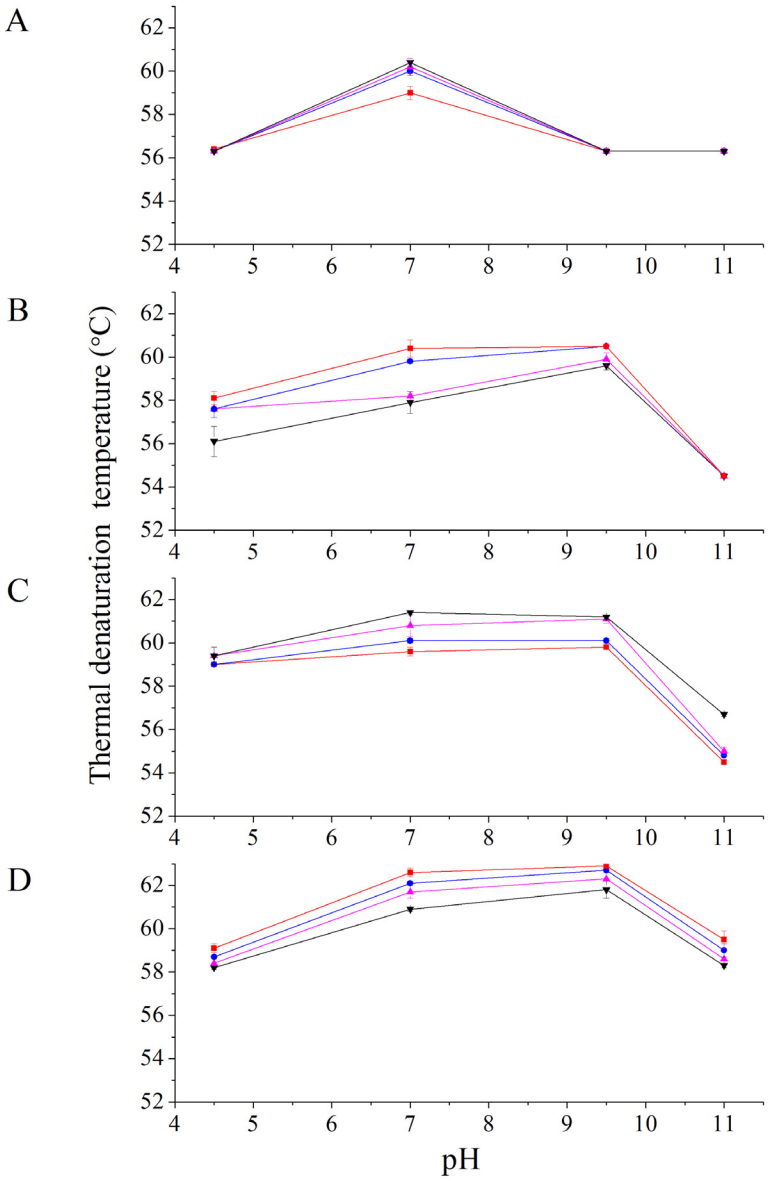


Figure 4.17 – The primary thermal denaturation temperatures of **(A)** DHDPS A207V, **(B)** pET M11 DHDPS A207V, **(C)** pET 151/D-TOPO DHDPS A207V and **(D)** cleaved pET 151/D-TOPO DHDPS A207V in 100 mM phosphate buffer at 0 mM NaCl (red), 40 mM NaCl (blue), 100 mM NaCl (pink) and 500 mM NaCl (black) plotted as a function of pH. The mean thermal denaturation temperatures plotted were calculated from derivations of the fluorescence data. The error bars represent the SEM.

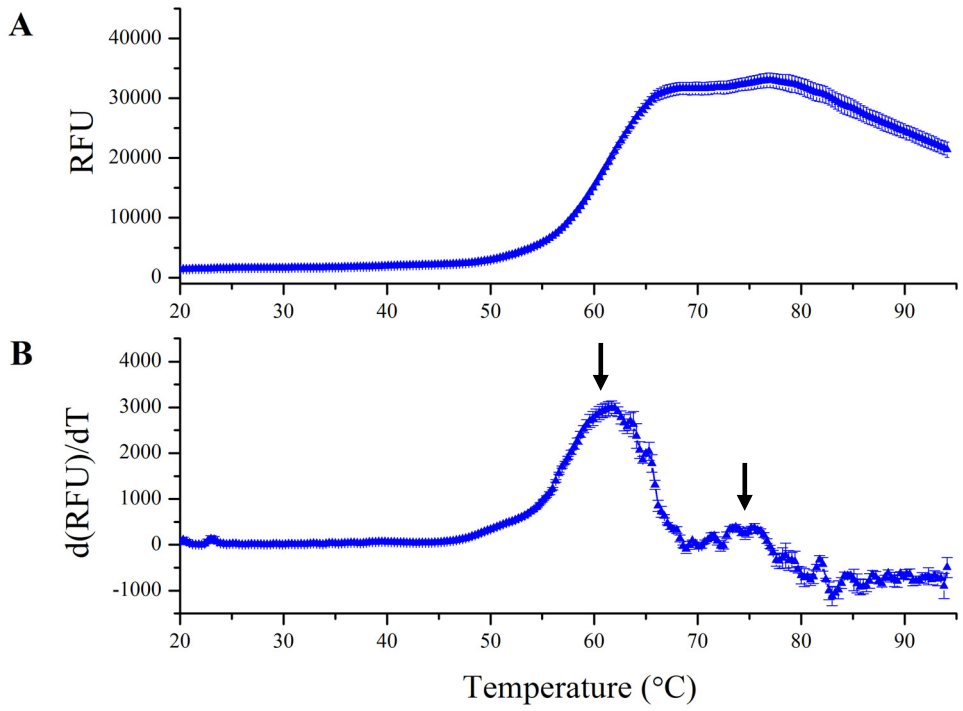


Figure 4.18 – The biphasic thermal denaturation of pET 151/D-TOPO DHDPS Q90L (blue) in 100 mM phosphate buffer (pH 7, containing 0 mM NaCl) as monitored by SYPRO Orange fluorescence. **(A)** The increase in fluorescence, as monitored by the BioRad IQ5. **(B)** The derivatives of the data presented in figure 4.14. The peaks indicated by the black arrows indicate the temperatures at which the protein is unfolding.

DSF demonstrated that the introduction of the Q90L and A207V point mutations and/or the attachment of a polyhistidine tag to *E. coli* DHDPS caused changes (although often non-significant) in the thermal denaturation of the proteins. The thermal denaturation temperature of DHDPS Q90L using this technique was not quantifiable due to high initial fluorescence. The presence of the valine at position 207 caused a general decrease in thermal denaturation temperature. The addition of the pET M11 polyhistidine tag to DHDPS caused an increase of the thermal denaturation temperature of DHDPS Q90L compared to wild-type DHDPS. The addition of the pET M11 polyhistidine-tagged to A207V caused no significant difference compared to DHDPS A207V. The pET 151/D-TOPO polyhistidine tag caused the thermal denaturation temperature of DHDPS Q90L to decrease compared to pET M11 DHDPS Q90L. In the case of DHDPS A207V, the pET 151/D-TOPO polyhistidine tag caused an increase in thermal denaturation temperature. The residues remaining following the cleavage of the pET 151/D-TOPO polyhistidine tag generally decreased the thermal denaturation temperature compared to pET 151/D-TOPO DHDPS Q90L. The cleaved pET

151/D-TOPO polyhistidine-tagged A207V demonstrated a generally higher thermal denaturation than DHDPS A207V, pET M11 DHDPS A207V and pET 151/D-TOPO DHDPS A207V.

DSF demonstrated that the amino acid substitutions predicted by the algorithm, Zyggregator altered the thermal stability of DHDPS. The presence of the polyhistidine tags and the residues remaining following the cleavage of the pET 151/D-TOPO polyhistidine tag, further altered the thermal denaturation temperatures of DHDPS Q90L and A207V but not in a systematic fashion. In order to corroborate the thermal denaturation assay results, CD spectroscopy was employed.

4.3.4.6 CIRCULAR DICHROISM SPECTROSCOPY

The secondary structures of DHDPS Q90L, pET M11 DHDPS Q90L, pET 151/D-TOPO DHDPS and cleaved pET 151/D-TOPO DHDPS were assessed using CD spectroscopy and compared to that of wild type DHDPS. Figure 4.19 shows that the wavelength spectra of the variants are generally similar, indicating that neither the presence of the Q90L and A207V amino acid substitutions nor the combinatorial effect of polyhistidine tags or amino acids had any effect on the secondary structure of *E. coli* DHDPS in solution. All the spectra have double minima at 208 nm and 222 nm which are the characteristic features of $(\alpha/\beta)_8$ barrel structure. Despite exhibiting double minima, the spectra obtained for DHDPS Q90L did demonstrate some variance from the wild-type enzyme. It is possible that these aberrant secondary structures lead to the un-quantifiable DSF results obtained from DHDPS Q90L and may affect the formation of amorphous and β-sheet-specific aggregates. This will be discussed in sections 4.3.2.7 and 4.3.2.8. Such a change in the biophysical properties of DHDPS Q90L may have also been the reason crystals were not obtained for this variant (see section 4.3.4.2).

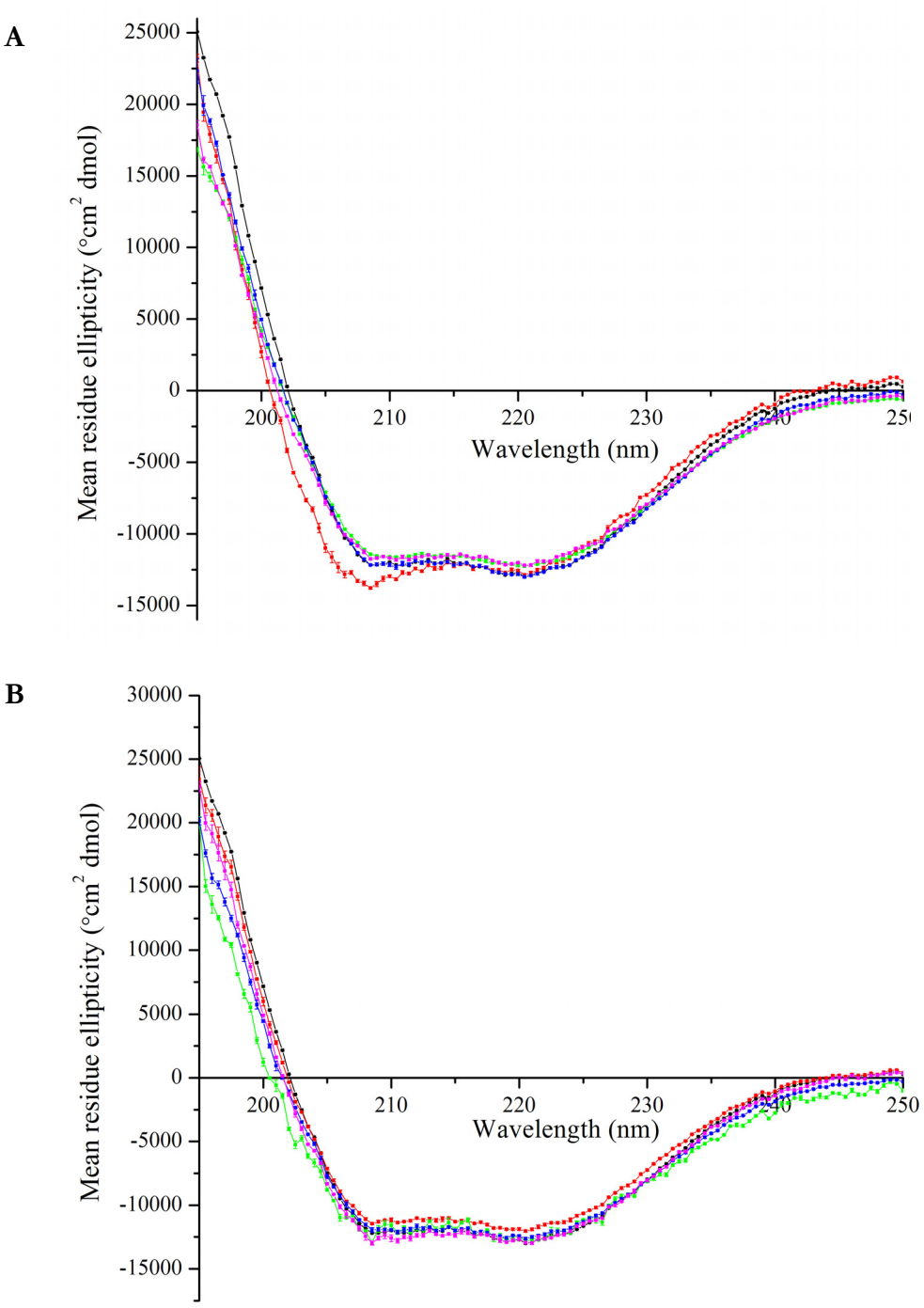


Figure 4.19 – The results of wavelength scans of **(A)** wild-type DHDPS (black), DHDPS Q90L (red), pET M11 DHDPS Q90L (green), pET 151/D-TOPO DHDPS Q90L (blue) and cleaved pET 151/D-TOPO DHDPS Q90L (pink). **(B)** Wild-type DHDPS (black) DHDPS A207V (red), pET M11 DHDPS A207V (green), pET 151/D-TOPO DHDPS A207V (blue) and cleaved pET 151/D-TOPO DHDPS A207V (pink). All assays were performed in 20 mM Tris (pH 7, 40 mM NaCl). The plot shows mean residue ellipticity vs. wavelength (N = 3) ±SEM.

The thermal denaturation temperatures of DHDPS Q90L, DHDPS A207V and their polyhistidine-tagged species were established by monitoring the change in intensity at 220 nm as a function of temperature. All demonstrated similar denaturation profiles to that of wild-type DHDPS and are presented in their entirety in appendix 3, section A3.3. The overall trends observed in the CD melts mirrored the outcome of the DSF experiments (table 4.6) although the thermal denaturation temperature determined by CD is consistently lower than that obtained by DSF, which was discussed in chapter 2, section 2.3.2.6. The thermal denaturation temperatures for pET M11 DHDPS A207V demonstrated large variability between CD spectroscopy and DSF. This disparity may have been a reflection of an unstable structure, a hypothesis corroborated by poor purification yields (section 4.3.3), kinetic analyses which demonstrated a reduced catalytic ability (section 4.3.4.1) and analytical gel permeation liquid chromatography which demonstrated some anomalous results (section 4.3.4.3).

Variant	Thermal denaturation temperature as per CD (°C)	Thermal denaturation temperatures as per DSF (°C)
Wild-type DHDPS	56.9	60.9 ± 0.3
DHDPS Q90L	54.4	-
pET M11 DHDPS Q90L	58.8	63.1 ± 0.1
pET 151/D-TOPO DHDPS Q90L	59.9	61.7 ± 0.0
Cleaved pET 151/D-TOPO DHDPS Q90L	59.6	61.2 ± 0.1
DHDPS A207V	52.1	60.0 ± 0.2
pET M11 DHDPS A207V	50.3	58.2 ± 0.1
pET 151/D-TOPO DHDPS A207V	59.2	60.1 ± 0.1
Cleaved pET 151/D-TOPO DHDPS A207V	55.2	61.7 ± 0.3

Table 4.6 – The thermal denaturation temperatures of DHDPS Q90L and DHDPS A207V and their polyhistidine-tagged variants in 20 mM phosphate buffer (pH 7, 40 mM NaCl) as monitored by CD spectroscopy and DSF. There are no errors on the CD data as discussed in chapter 3, section 3.3.3.5 and the DSF data is that from the primary thermal denaturation temperature.

The thermal denaturation temperatures determined by the CD melt experiments confirmed that the presence of the amino acid substitutions at positions 90 and 207 affect the thermal denaturation temperatures of DHDPS, causing a decrease for both non-tagged variants (see table 4.6). The addition of the polyhistidine tags to the variants raised the thermal denaturation temperatures, possibly reflecting the increased size of the tagged proteins. The pET M11 polyhistidine tag appears to have stabilised the protein containing the Q90L amino acid substitution, whereas it has significantly destabilised DHDPS A207V. This result is perhaps reflected in the difficulty purifying the pET M11 polyhistidine-tagged *E. coli* DHDPS A207V (see section 4.3.3).

The results of the CD spectroscopy have ascertained that the secondary structures of DHDPS, pET M11 DHDPS Q90L, pET 151/D-TOPO DHDPS Q90L, cleaved pET 151/D-TOPO DHDPS Q90L, DHDPS A207V, pET M11 DHDPS A207V, pET 151/D-TOPO DHDPS A207V, and cleaved pET 151/D-TOPO DHDPS A207V are similar to that of wild-type enzyme. DHDPS Q90L exhibited a different CD spectrum, potentially explaining the inability to obtain crystals or quantifiable DSF data. The thermal denaturation temperatures supported the results of the DSF; the presence of polyhistidine motifs typically increased the thermal denaturation temperatures of DHDPS Q90L and A207V except for in the case of pET M11 DHDPS A207V, which exhibited a lower thermal denaturation temperature than DHDPS A207V. The effect of the altered thermal stabilities of the variants with regard to amorphous and β-sheet-specific aggregation is discussed in the following sections (sections 4.3.4.7 and 4.3.4.8).

4.3.4.7 AMORPHOUS AGGREGATION

The aggregation propensities of the DHDPS Q90L and DHDPS A207 variants were carried out as described in 2, section 2.3.2.7. Boltzmann analysis was used to assess the half life of aggregation ($Agg_{½}$) and, as described for other variants, the analysis was restricted to the results for pH 7 and pH 9.5 (and in some cases pH 4.5, 500 mM NaCl and pH 11, 500 mM NaCl) due to rapid aggregation or no increase in light scattering under other conditions. The maximum absorbance was measured during the initial 90 minutes of the assay. The results are summarised below in figures 4.20 – 4.23 and in appendix 3, section A3.4. Statistical analysis of the $Agg_{½}$ and maximum aggregation was carried out using the technique described in chapter 3, section 3.3.3.6, however, only the results at pH 7 were used in this statistical analysis of $Agg_{½}$ because the data sets were not complete for other pH values. Significance is presented to a P value of < 0.01. The analysis of deviance for $Agg_{½}$ is presented in appendix 3, table A3.9. The analysis of deviance for maximum absorbance is presented in appendix 3, table A3.10.

The results of this assay suggested that the presence of the amino acid substitution at position 90 generally decreased the aggregation propensity of DHDPS with regard to the maximum light scattering (see figure 4.20). There were some conditions under which the aggregation propensity was increased, particularly pH 4.5 and under high salt conditions at pH 7 and pH 9.5. When combined with the polyhistidine motifs, the speed of aggregation generally increased significantly compared to

DHDPS Q90L. However, there were exceptions to this, with pET 151/D-TOPO DHDPS Q90L and cleaved pET 151/D-TOPO DHDPS Q90L undergoing slower aggregation than both wild-type and DHDPS Q90L at pH 4.5. For both pET 151/D-TOPO DHDPS Q90L and cleaved pET 151/D-TOPO DHDPS Q90L, the maximum light scattering was significantly reduced across all conditions. This may be a reflection of different aggregate sizes as discussed in chapter 3, section 3.3.3.6. As was observed for the tagged wild-type protein, the addition of the pET M11 polyhistidine tag to DHDPS Q90L caused significant increases in both the speed of aggregation and the maximum light scattering compared to DHDPS Q90L. These results corroborate those published by Dümmler *et al.* [31] who found that the pET M11 polyhistidine tag caused a significant decrease in protein solubility and stability. Although, from the comparison of DHDPS Q90L and its polyhistidine-tagged variants with wild-type DHDPS, it can be concluded that the amino acid substitution at position 90 generally decreased the aggregation propensity of DHDPS. The effect of the polyhistidine tags and the residues remaining following the cleavage of the pET 151/D-TOPO polyhistidine tag had the most impact, increasing the speed of aggregation and, in the case of the pET M11 polyhistidine tag, significantly increasing the maximum light scattering.

Under all conditions, DHDPS A207V and its polyhistidine-tagged variants demonstrated a longer aggregation half life and decreased maximum light scattering compared to wild-type (figure 4.22). This indicates that DHDPS A207V is less likely to form amorphous aggregate than wild-type. This result contrasts with the results obtained for DHDPS Q90L which exhibited a higher tendency to aggregate than wild-type enzyme under some conditions. The combinatorial affects of A207V and the polyhistidine motifs resulted in an increase in the $Agg_{½}$ except at pH 4.5, pH 7, 100 mM NaCl and pH 7 500 mM NaCl, for which all the polyhistidine-tagged variants of DHDPS A207V aggregated more slowly than both DHDPS A207V and wild-type DHDPS. The aggregation propensity of cleaved pET 151/D-TOPO DHDPS A207V was unable to be assessed due to the large quantity of protein required for the assay, the low purification yields (see section 4.3.3) and the prioritisation of the ThT assays (section 4.3.4.8). However, in light of the results obtained for the cleaved wild-type (see figure 4.22 and chapter 3, section 3.3.3.7) it is reasonable to assume that the residues remaining following cleavage would decrease or cause no significant change in the aggregation propensity of DHDPS A207V compared to pET 151/D-TOPO DHDPS A207V.

The effects of both salt and pH on the maximum light scattering of all the DHDPS A207V variants were similar to that observed for wild-type DHDPS and are shown in figure 4.23. Maximum light

scattering was observed at pH 4.5 followed by pH 7 and pH 9.5. Very low readings were obtained at pH 2 and pH 11, except for in the presence of 500 mM NaCl for pET 151/D-TOPO DHDPS, which exhibited increased maximum absorbance at both high and low pH.

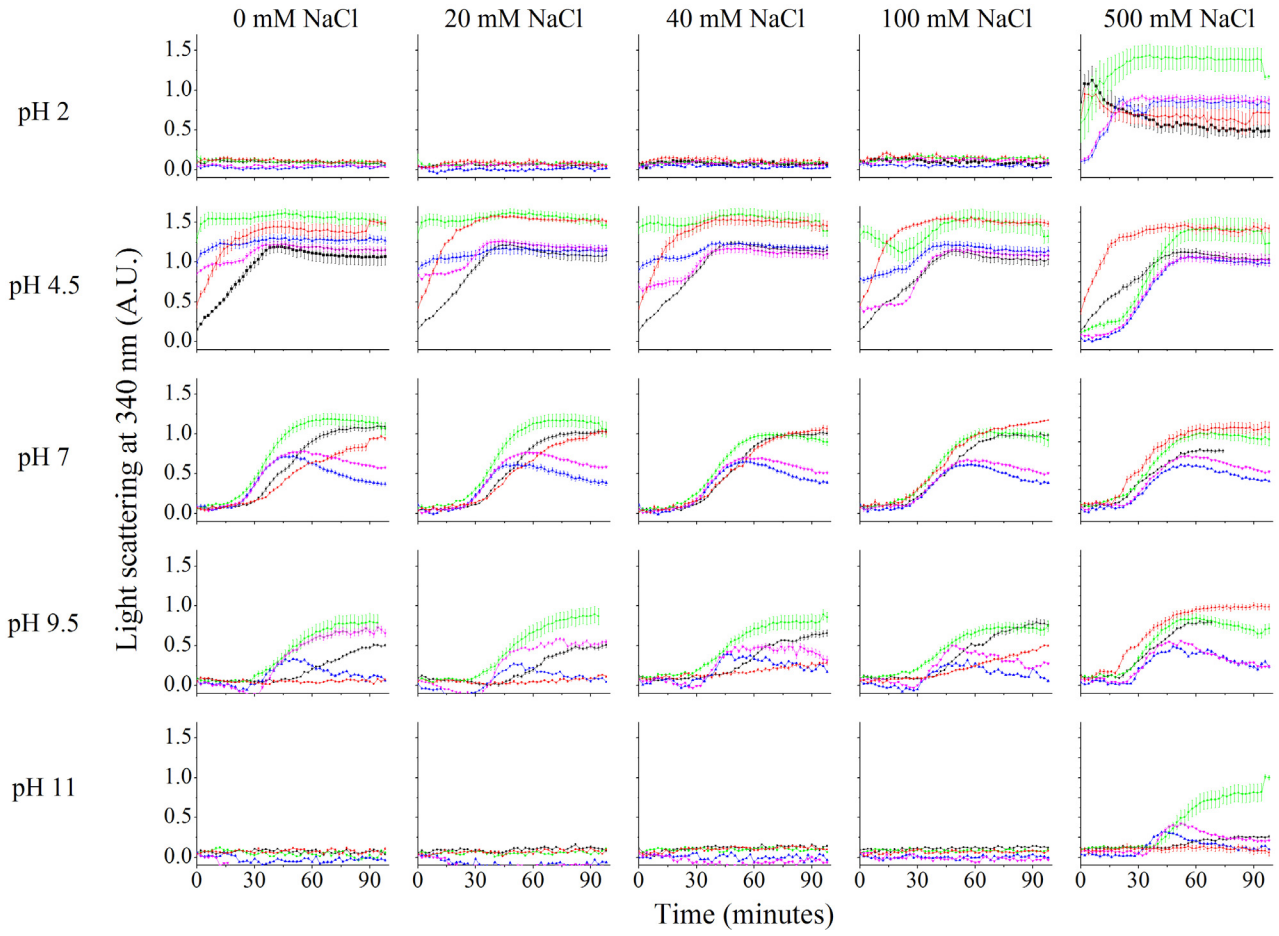


Figure 4.20 – The aggregation propensity of wild-type *E. coli* DHDPS (black), *E. coli* DHDPS Q90L (red), pET M11 *E. coli* DHDPS Q90L (green), pET 151/D-TOPO *E. coli* DHDPS Q90L (blue) and cleaved pET 151/D-TOPO *E. coli* DHDPS Q90L (pink) in 100 mM phosphate buffer (pH 7, 40 mM NaCl) in 100 mM phosphate buffer at 60 ºC. Where possible a sigmoid (Boltzmann) function was used for the line of best fit (> 98% confidence interval). Once the peak absorbance was reached, the absorbance was inconsistent due to the sedimentation of the protein aggregate [32]. The data plotted are the mean values of six replicates ± SEM. The aggregation half life ($Agg_{½}$) was determined by identifying the midpoint of the sigmoid curve as previously described [33]. The maximum absorbance was the highest reading over the initial 90 minutes of the assay.

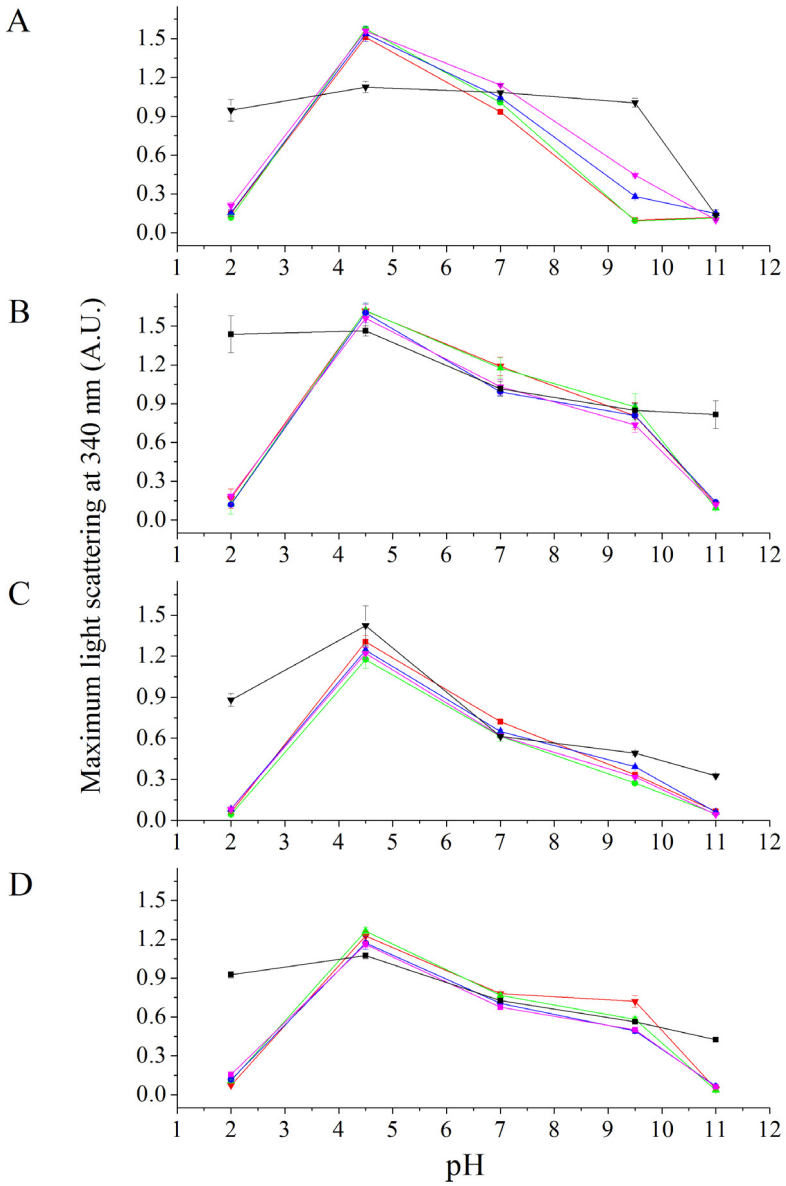


Figure 4.21 – The maximum absorbance of **(A)** DHDPS Q90L, **(B)** pET M11 DHDPS Q90L, **(C)** pET 151/D-TOPO DHDPS Q90L and **(D)** cleaved pET 151/D-TOPO DHDPS Q90L in 100 mM phosphate buffer, 0 mM NaCl (red), 20 mM NaCl (green), 40 mM NaCl (blue), 100 mM NaCl (pink) and 500 mM NaCl (black). The data plotted are the maximum absorbance units reached over the initial 90 minutes of the assay ± SEM and are plotted as a function of pH. These results are also presented in appendix 3, section A3.4.

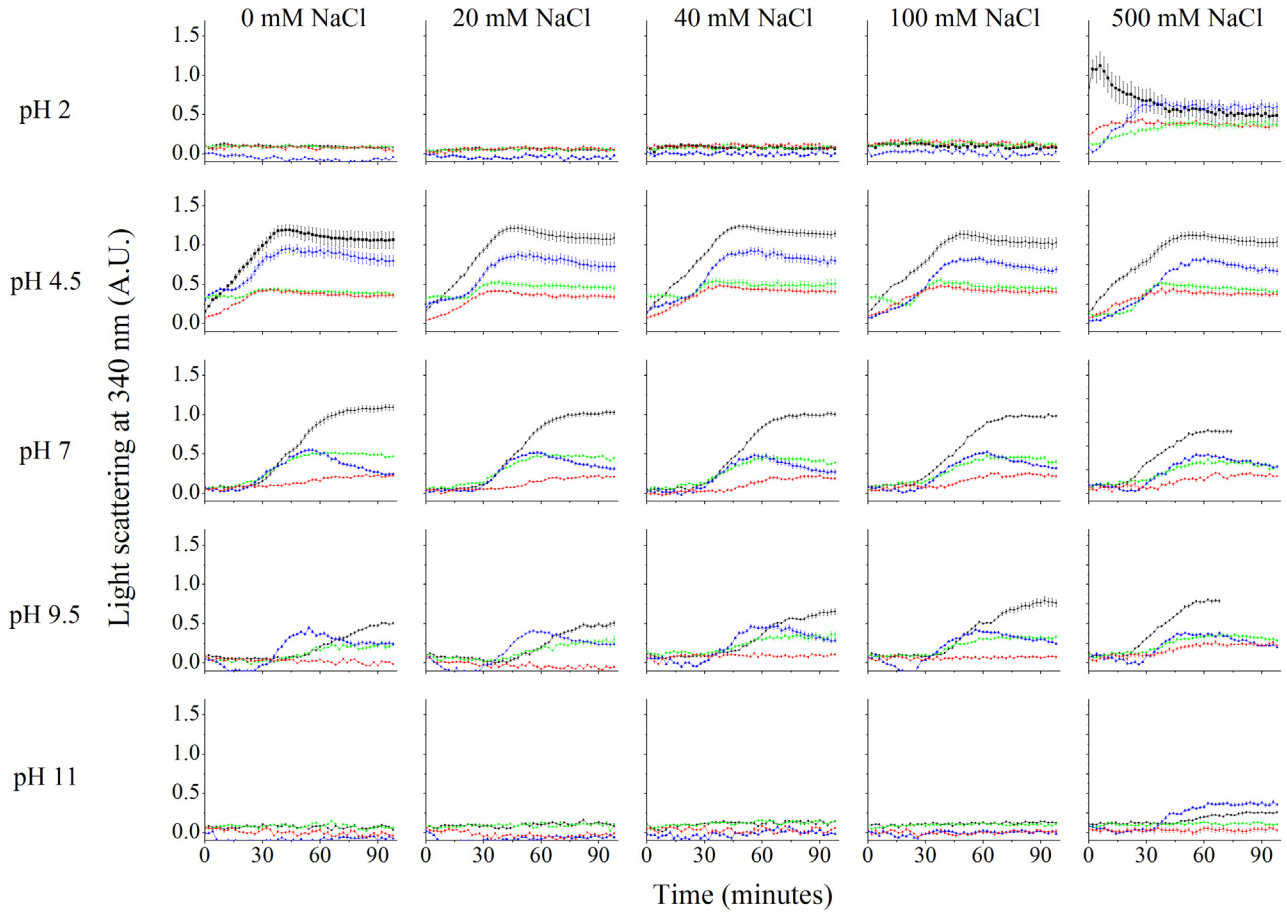


Figure 4.22 – The aggregation propensity of wild-type *E. coli* DHDPS (black), *E. coli* DHDPS A207V (red), pET M11 polyhistidine-tagged *E. coli* DHDPS A207V (green), and pET 151/D-TOPO polyhistidine-tagged *E. coli* DHDPS A207V (blue) in 100 mM phosphate buffer at 60 ºC. Where possible a sigmoid (Boltzmann) function was used for the line of best fit (> 98% confidence interval). Once the peak absorbance was reached, the absorbance was inconsistent due to the sedimentation of the protein aggregate [32]. The data plotted are the mean values and SEM (n = 6) ± SEM. The aggregation half-life ($Agg_{½}$) was determined by identifying the midpoint of the sigmoid curve as previously described [33]. The maximum absorbance was the highest reading over the initial 90 minutes of the assay.

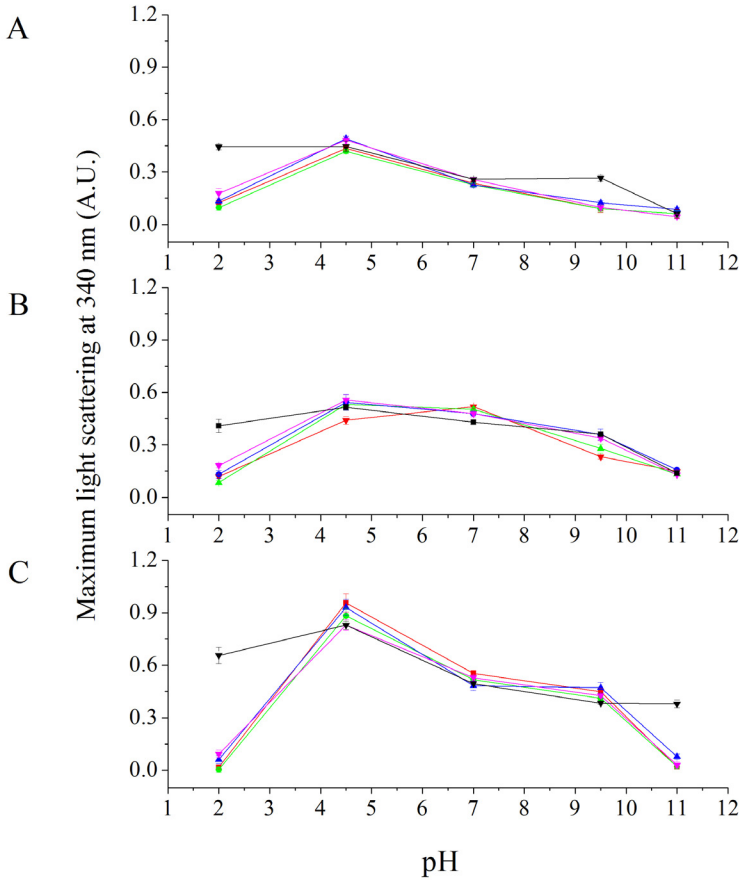


Figure 4.23 – The maximum absorbance of **(A)** DHDPS A207V, **(B)** pET M11 DHDPS A207V, **(C)** pET 151/D-TOPO DHDPS A207V and **(D)** cleaved pET 151/D-TOPO DHDPS A207V in 100 mM phosphate buffer, 0 mM NaCl (red), 20 mM NaCl (green), 40 mM NaCl (blue), 100 mM NaCl (pink) and 500 mM NaCl (black). The data plotted are the maximum absorbance units reached over the initial 90 minutes of the assay ±SEM and are plotted as a function of pH. These results are also presented in appendix 3, section A3.4.

The effect of salt concentration on the $Agg_{½}$ of DHDPS Q90L and A207V was similar to that observed for wild-type DHDPS. The results of the assay demonstrate that both DHDPS Q90L and A207V undergo the characteristic decrease in $Agg_{½}$ with increasing salt concentration at pH 9.5 (see appendix A3, section A3.4). As discussed in chapter 3, section 3.3.3.6, the presence of the polyhistidine tags appears to have reduced the effect of the solvating power of NaCl on DHDPS irrespective of the presence of the amino acid substitutions at positions 90 and 207 [28]. The effect of pH on the $Agg_{½}$ of DHDPS Q90L and A207V was similar to wild-type which underwent more rapid aggregation at pH 4.5 followed by pH 7 and pH 9.5. As was observed in the case of the polyhistidine-tagged variants of wild-type, the effects of pH on the speed of aggregation were reduced for all the polyhistidine-tagged variants of DHDPS Q90L and A207V.

The results of this assay indicated that the presence of pET M11 polyhistidine tag, pET 151/D-TOPO polyhistidine tag and the residues remaining following cleavage of the pET 151/D-TOPO polyhistidine tag play a major role in determining the aggregation propensity of Q90L and A207V, supporting the conclusions drawn in chapter 3. The algorithmically predicted amino acid substitution at position 90 generally increased the propensity of DHDPS to form amorphous aggregate. However, the presence of a valine in place of an alanine at position 207 generally decreased the aggregation propensity of DHDPS. As the algorithm is designed to predict changes in β-sheet-specific aggregation, ThT assays were performed and are described below (section 4.3.4.8).

4.3.4.8 β-SHEET-SPECIFIC AGGREGATION

The propensities of DHDPS Q90L, A207V and their polyhistidine-tagged variants to form β-sheet-rich aggregate were screened in 100 mM phosphate buffer at pH 2, pH 4.5, pH 7, pH 9.5 and pH 11 each containing 0 mM, 40 mM, and 500 mM NaCl and in the presence of 1.6 mM ThT. Each assay was carried out as described chapter 8, section 8.6.6, and the results were compared to those obtained for wild-type DHDPS. Statistical analysis was carried out using the techniques described in chapter 3, section 3.3.3.7. Significance is presented to a P value of < 0.01. The analysis of deviance for $\beta\text{-}agg_{½}$ is presented in appendix 3, table A3.13. The analysis of deviance for maximum absorbance is presented in appendix 3, table A3.14.

As discussed in chapter 2, section 2.3.2.8, comparison between conditions could not be made due to the effect of salt on the kinetics of β-sheet-specific aggregation and binding of ThT, and the effect of pH on the affinity of ThT for β-sheet structures. However, under each condition, the effects of the amino acid substitutions and the combinatorial effects of the polyhistidine tags were compared to the results obtained for the wild-type protein. This comparative analysis revealed that the amino acid substitution at position 90 caused a decrease in the ThT fluorescence. In combination with the polyhistidine tags, the Q90L substitution resulted in a significant increase in ThT fluorescence across all conditions (see figure 4.24). This result is similar to the pattern observed for the polyhistidine variants of wild-type DHDPS. However, the combinatorial effect of the polyhistidine tags and the amino acid substitution at position 90 resulted in a significantly higher maximum fluorescence than the corresponding polyhistidine-tagged variants of wild-type DHDPS. The pET M11 DHDPS Q90L resulted in the highest fluorescence followed by the pET 151/D-TOPO DHDPS Q90L. The amino acid substitution at position 90 in combination with the residues remaining following cleavage of the

pET 151/D-TOPO polyhistidine tag resulted in a significantly higher fluorescence than observed for the cleaved pET 151/D-TOPO DHDPS species or wild-type DHDPS. This indicated that the combinatorial effect of the polyhistidine tags or the residues remaining following cleavage of the pET 151/D-TOPO polyhistidine tag with the amino acid substitution at position 90 resulted in a DHDPS variant which was more inclined to form β-sheet-specific aggregate than the wild-type protein.

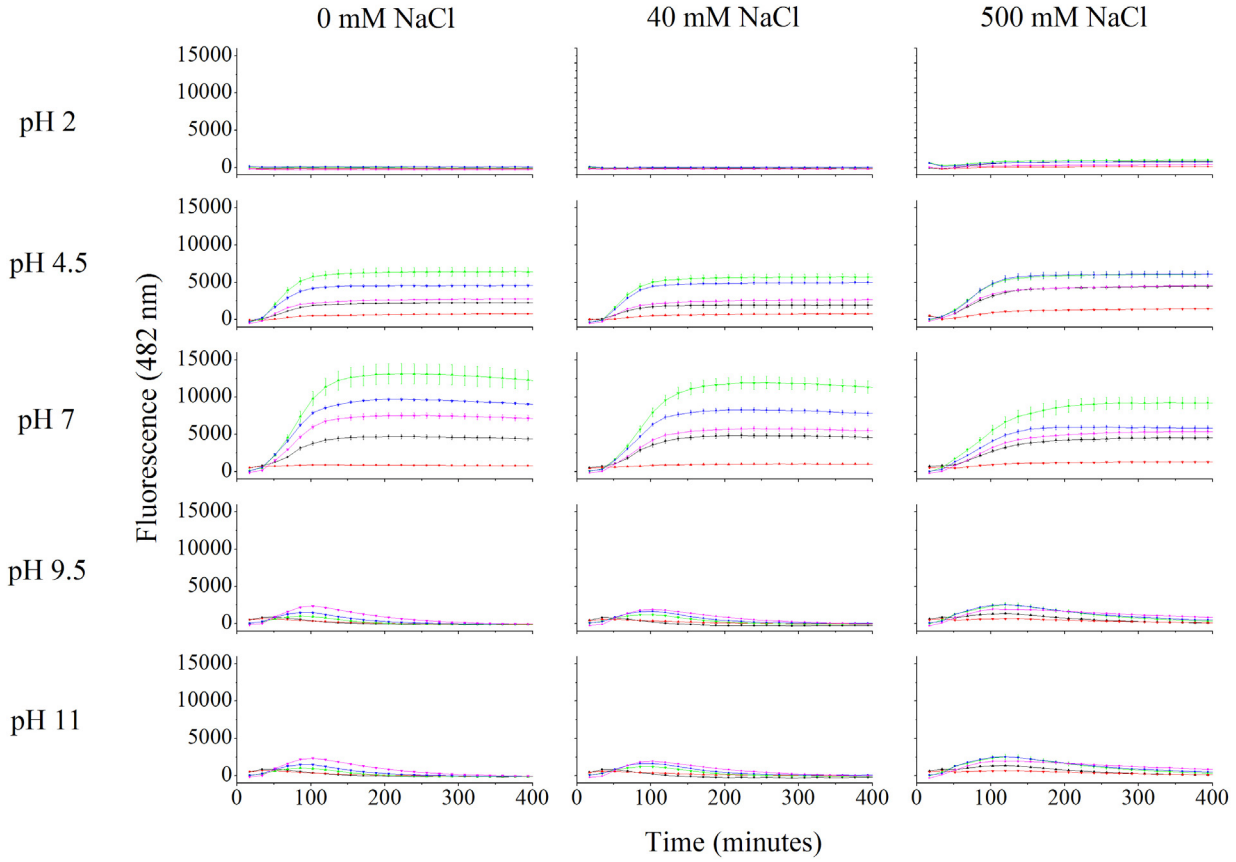


Figure 4.24 – The thioflavin T induced fluorescence of wild-type DHDPS (black), DHDPS Q90L (red), pET M11 DHDPS Q90L (green), pET 151/D-TOPO DHDPS Q90L (blue) and cleaved pET 151/D-TOPO DHDPS Q90L (pink) in 100 mM phosphate buffer at pH 2, pH 4.5, pH 9.5 and pH 11 containing 0 mM NaCl, 40 mM NaCl and 500 mM NaCl. Where possible a sigmoid (Boltzmann) distribution was used for the line of best fit (> 98% confidence interval). Note that only the data up until the linear phase were included in the analysis of $\beta\text{-}agg_{½}$. Once the peak absorbance was reached the absorbance was inconsistent due to the sedimentation of the protein aggregate. The maximum fluorescence was the highest reading over the initial 300 minutes of the assay. The plotted data are the mean of six replicates ± SEM.

The amino acid substitution at position 207, like Q90L, reduced the maximum fluorescence significantly. However, unlike the results obtained for the polyhistidine-tagged variants of Q90L, the presence of the polyhistidine tags in combination the amino acid substitution A207V, did not result in vastly increased fluorescence in most cases (figure 4.25). The only DHDPS A207V variant exhibiting any increase in fluorescence was pET 151/D-TOPO DHDPS A207V. The pET M11 polyhistidine tag induced fluorescence increase observed for both wild-type DHDPS (see chapter 3, section 3.3.3.7) and DHDPS Q90L (see figure 4.24) appeared to have been reduced by the presence of the valine at position 207 in place of alanine.

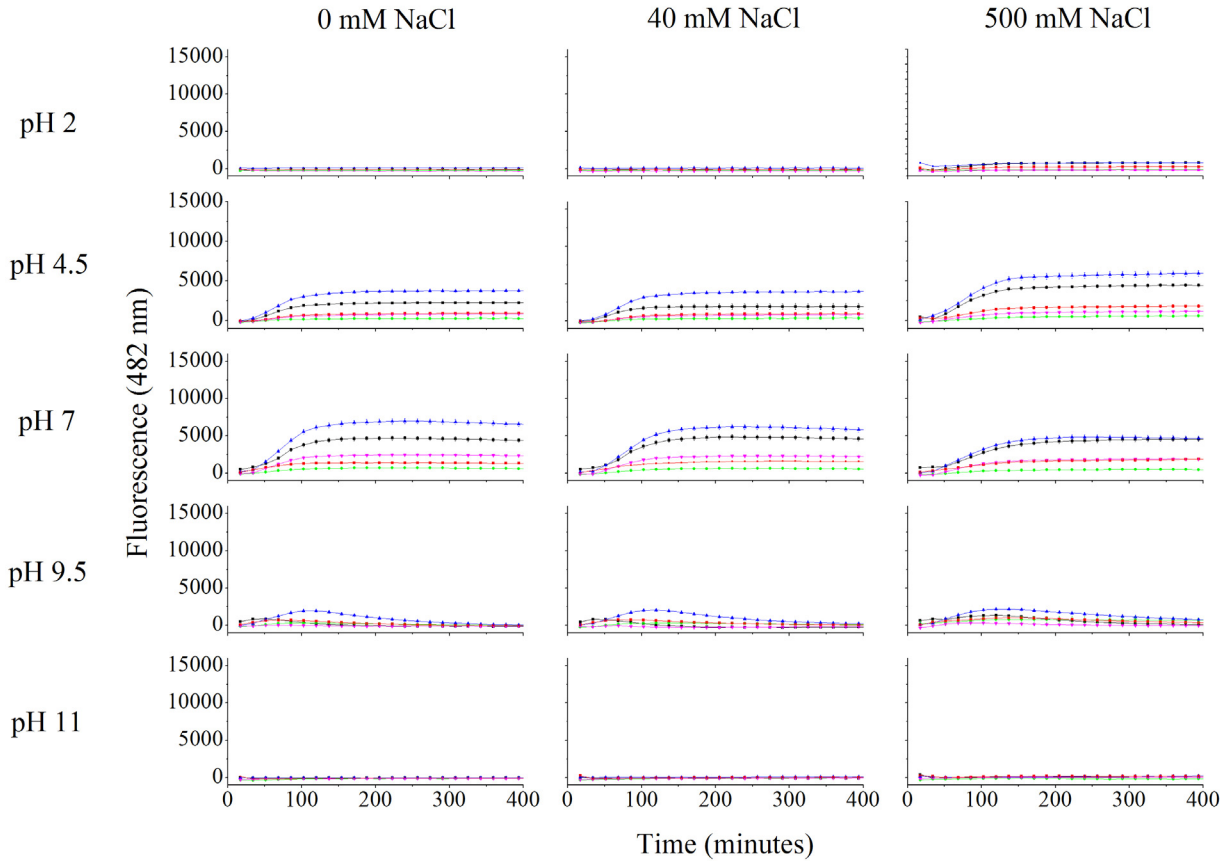


Figure 4.25 – The thioflavin T induced fluorescence of wild-type DHDPS (black), DHDPS A207V (red), pET M11 DHDPS A207V (green), pET 151/D-TOPO DHDPS A207V (blue) and cleaved pET 151/D-TOPO DHDPS A207V (pink) in 100 mM phosphate buffer at pH 2, pH 4.5, pH 9.5 and pH 11 containing 0 mM NaCl, 40 mM NaCl and 500 mM NaCl. Where possible sigmoid (Boltzmann) distributions were used for the line of best fit (> 98% confidence interval). Note that only the data up until the linear phase were included in the analysis of β-$agg_{½}$. Once the peak absorbance was reached the absorbance was inconsistent due to the sedimentation of the protein aggregate. The maximum fluorescence was the highest reading over the initial 300 minutes of the assay. The plotted data are the mean of six replicates ± SEM.

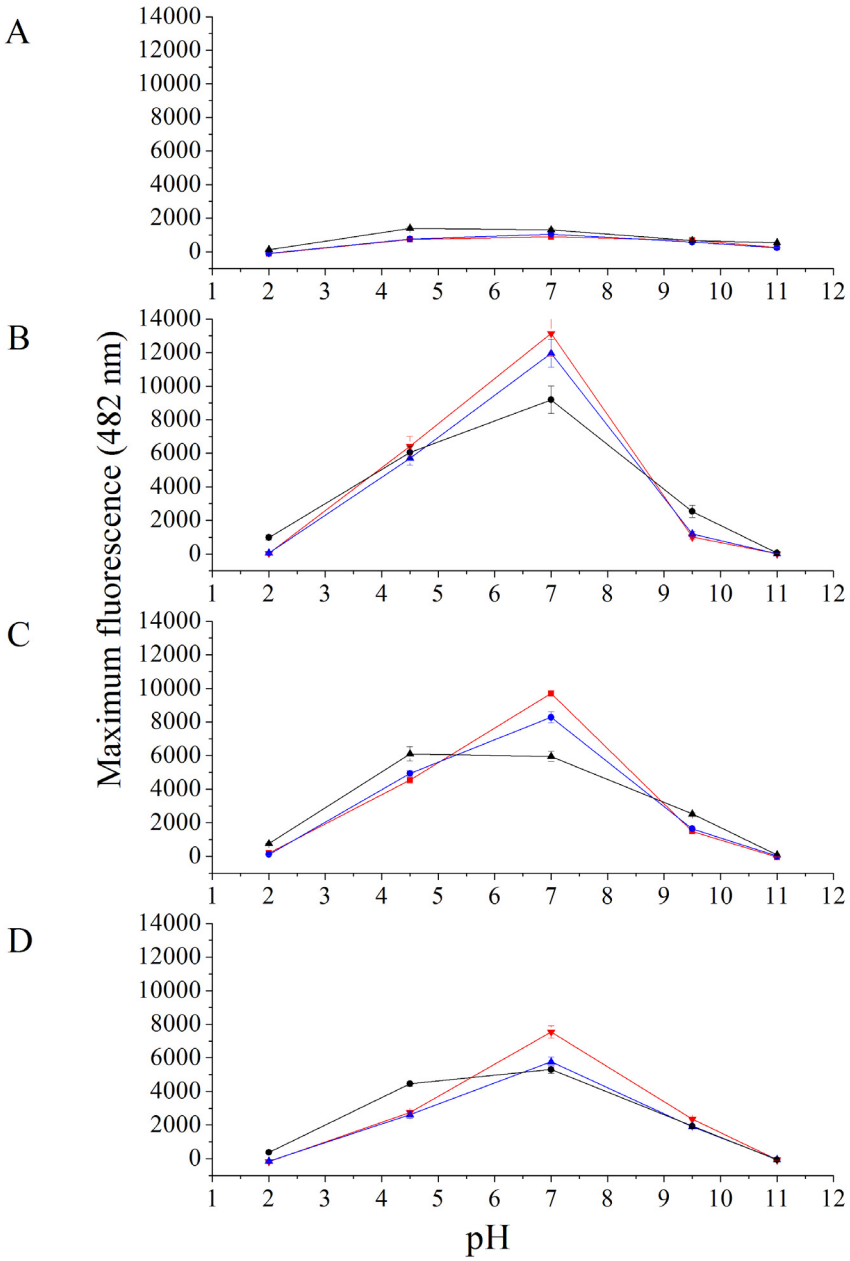


Figure 4.26 – The maximum fluorescence of **(A)** DHDPS Q90L, **(B)** pET M11 DHDPS Q90L, **(C)** pET 151/D-TOPO DHDPS Q90L and **(D)** cleaved pET 151/D-TOPO DHDPS Q90L in 100 mM phosphate buffer, 0 mM NaCl (red), 40 mM NaCl (blue) and 500 mM NaCl (black). The data plotted are the mean maximum fluorescence reached over the initial 300 minutes of the assay ± SEM and are plotted as a function of pH.

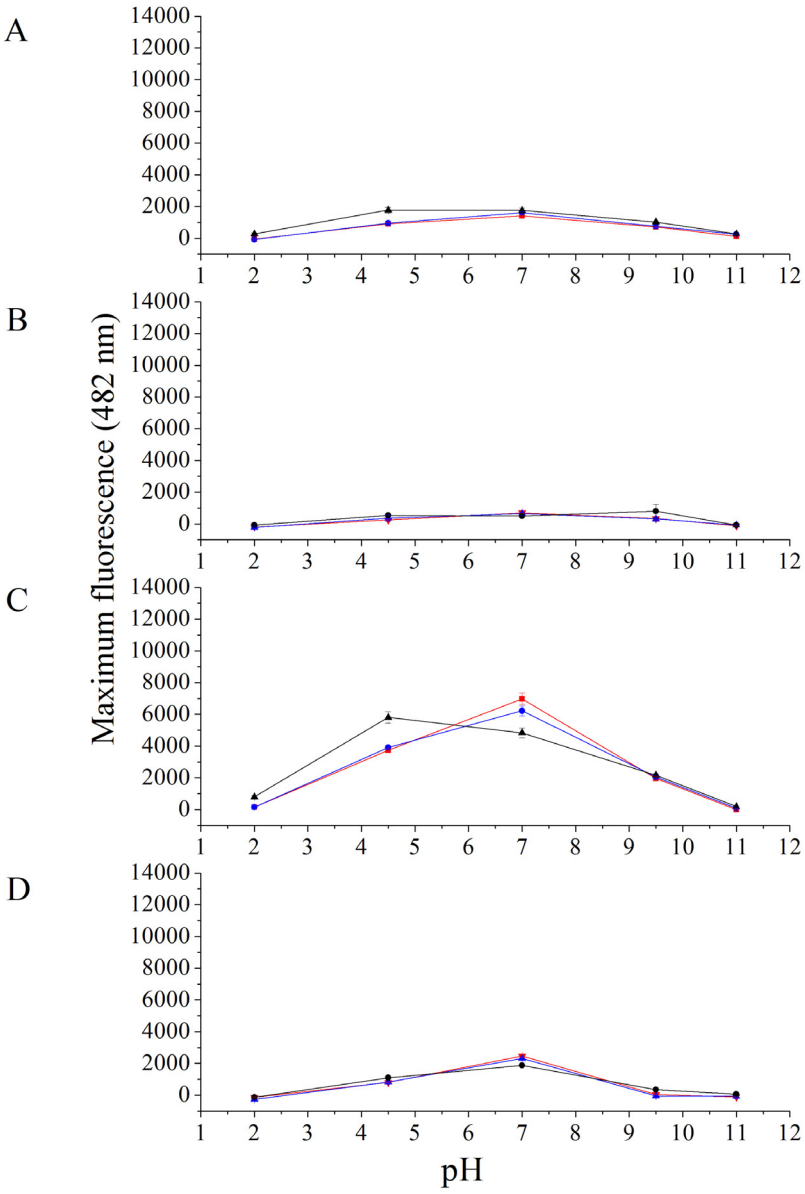


Figure 4.27 – The maximum fluorescence of **(A)** DHDPS A207V, **(B)** pET M11 DHDPS A207V, **(C)** pET 151/D-TOPO DHDPS A207V and **(D)** cleaved pET 151/D-TOPO DHDPS A207V in 100 mM phosphate buffer, 0 mM NaCl (red), 40 mM NaCl (blue) and 500 mM NaCl (black). The data plotted are the mean maximum fluorescence reached over the initial 300 minutes of the assay ± SEM and are plotted as a function of pH.

The increased maximum fluorescence observed in the case of the DHDPS Q90L polyhistidine-tagged variants and pET 151/D-TOPO DHDPS A207V suggested that these species were more likely to form β-sheet-specific aggregate than DHDPS Q90L or the wild-type protein [34-38]. As fluorescence of ThT in the presence of β-sheet-specific aggregate is considered indicative of amyloid

fibril formation, further attempts to confirm the presence of fibril-like structures were pursued through both TEM and X-ray fibre diffraction.

The decreased maximum fluorescence observed for DHDPS A207V, pET M11 DHDPS A207V and cleaved pET 151/D-TOPO DHDPS A207V suggest that these variants were less prone to forming β-sheet-specific aggregate than wild-type DHDPS. Due to this, the confirmation of the formation of fibrillar structures was not pursued for these variants.

4.3.4.9 CONFIRMATION OF AMYLOID FIBRILS

Confirmation of the presence of amyloid fibrils by TEM was attempted in the DHDPS Q90L polyhistidine-tagged variants and pET 151/D-TOPO DHDPS A207V samples. These were prepared under the same conditions as those that exhibited the highest ThT induced fluorescence and as described in chapter 8, sections 8.8.1. No fibrillar structures were observed despite repeated attempts (figure 4.28). In addition to this, attempts to construct X-ray fibre diffraction stalks were unsuccessful (described in chapter 8, section 8.8.2). The failure to confirm the presence of amyloid fibrils suggests that the ThT results were not indicative of classical amyloid fibril formation but instead aggregates were β-sheet rich compared to the controls. This result was consistent with the fact that, thus far, no $(\alpha/\beta)_8$ barrel, including DHDPS, has been shown to form amyloid fibrils.

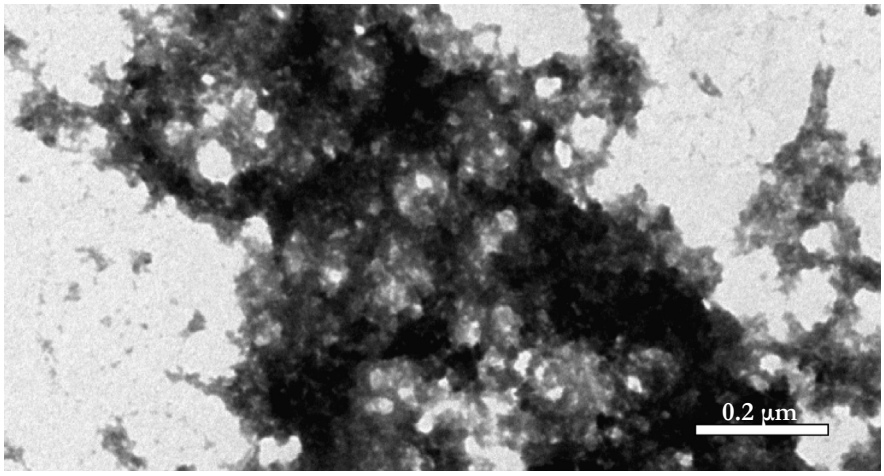


Figure 4.28 – TEM image of non-specific aggregate formed by pET M11 DHDPS Q90L in 40 mM NaCl at pH 7. This image was typical of all 5 images obtained and characteristic of the images obtained for all the variants screened using TEM.

4.5 SUMMARY AND CONCLUSIONS

The amino acid substitutions predicted by Zyggregator to increase the β-sheet-specific aggregation of *E. coli* DHDPS were characterised. The combinatorial effects of pET M11 polyhistidine tag, pET 151/D-TOPO polyhistidine tag and cleaved pET 151/D-TOPO polyhistidine tag were also assessed.

The analysis revealed that the amino acid substitutions at positions 90 and 207 affected the kinetics of the enzyme. A change in kinetic profile was also identified in the polyhistidine-tagged variants. A change in the kinetic profile of an enzyme can be indicative of a change in the structure of the protein. This possibility was investigated initially using crystallographic analysis. DHDPS A207V crystallised under the same conditions as the wild-type enzyme and the subsequently refined crystal structure demonstrated no significant change at any level of organisation. Crystals of DHDPS Q90L were not obtained, thus other techniques were used to establish the biophysical properties of this variant. Gel filtration confirmed that DHDPS Q90L, DHDPS A207V and their polyhistidine-tagged variants were, like wild-type, tetrameric in solution and AUC confirmed these results. CD spectroscopy confirmed that most of the variants had secondary structures similar to that of wild-type DHDPS. The only exception to this was DHDPS Q90L which yielded a CD spectrum that indicated that it was more disordered than the other variants. This was supported by the fact that crystals were not obtained for DHDPS Q90L and that the DSF data for this variant exhibited high initial fluorescence, indicative of a protein that had exposed hydrophobic patches prior to heating. The polyhistidine-tagged variants of DHDPS Q90L all exhibited higher thermal denaturation temperatures than the wild-type enzyme. DHDPS A207V and its polyhistidine-tagged variants all exhibited unchanged or lower thermal denaturation temperatures than the wild-type DHDPS. DSF also identified some of the variants as undergoing biphasic thermal denaturation, possibly indicating the presence of a second structure in the solution. These biphasic data indicated that for some of the DHDPS variants, a species with a lower thermal denaturation temperature than wild-type were present in solution. Other variants exhibited an apparent second thermal denaturation peak at a higher temperature than wild-type. Analytical gel permeation liquid chromatography or AUC did not identify any variability in the quaternary structure of the variants in solution, thus the lower thermal denaturation temperatures could be attributed to the thermally induced disruption of the structure. The higher thermal denaturation temperatures could be attributed to the presence of aggregates that

form upon heating. The results of the CD thermal denaturation experiments did not provide data regarding the presence of second species in solution; however, the thermal denaturation temperatures followed the general pattern of the DSF results.

Analysis of the propensity of DHDPS Q90L, DHDPS A207V and their polyhistidine variants to form amorphous and β-sheet-specific aggregate was carried out. The presence of the amino acid mutations at positions 90 and 207 did not result in DHDPS variants with an aggregation propensity that was different to that of the wild-type protein. The trends with regard to the aggregation propensity of the DHDPS Q90L and A207V variants were similar to those observed for the polyhistidine-tagged variants of wild-type DHDPS. The ThT fluorescence indicated that the amino acid substitutions at positions 90 and 207 decreased the propensity of non-tagged DHDPS to form β-sheet-specific aggregate. However, in the presence of the polyhistidine tags and the residues remaining following the cleavage of the pET 151/D-TOPO polyhistidine tag, the propensity of DHDPS Q90L to form β-sheet-specific aggregate was increased significantly compared to both wild-type DHDPS and its polyhistidine-tagged variants. No such effect was observed in the case of DHDPS A207V which only exhibited an increase in ThT fluorescence with regard to wild-type in the presence of the pET 151/D-TOPO polyhistidine tag, leading to the hypothesis that the presence of the valine at position 207 in place of the alanine rescued the protein from forming β-sheet-specific aggregate. Despite this increased propensity of the DHDPS Q90L polyhistidine-tagged variants to form β-sheet-specific aggregate, the presence of fibrillar structures by TEM or X-ray fibre diffraction was not achieved. These results indicate that the predictions of the algorithm, Zyggregator, (despite significantly increasing the hydrophilic/hydrophobic patterning in the amino acid sequence), were not accurate for DHDPS and that the presence of polyhistidine tags on DHDPS appears to be the major determinant of protein aggregation. The combinatorial affects of the amino acid substitutions and the extra-molecular residues were also unpredictable.

4.6 REFERENCES

1 Fernandez-Escamilla, A.-M., Rousseau, F., Schymkowitz, J. and Serrano, L. (2004) Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. Nature Biotechnology **22**, 1302-1306

2 Pawar, A. P., DuBay, K. F., Zurdo, J., Chiti, F., Vendruscolo, M. and Dobson, C. M. (2005) Prediction of "aggregation-prone" and "aggregation-susceptible" regions in proteins associated with neurodegenerative diseases. Journal of Molecular Biology **350**, 379-392

3 Guerois, R., Nielsen, J. E. and Serrano, L. (2002) Predicting changes in the stability of proteins and protein complexes: A study of more than 1000 mutations. Journal of Molecular Biology **320**, 369 - 387

4 Yang, Z. R., Thomson, R., McNeil, P. and Esnouf, R. M. (2005) RONN: The bio-basis function neural network technique applied to the detection of natively disordered regions in proteins. Bioinformatics **21**, 3369 - 3376

5 Parthiban, V., Gromiha, M. M. and Schomburg, D. (2006) CUPSAT: prediction of protein stability upon point mutation. Nucleic Acids Research **34**, W239 - W242

6 Capriotti, E., Fariselli, P. and Casadio, R. (2005) I-Mutant2.0: Predicting stability changes upon mutation form the protein sequence or structure. Nucleic Acids Research **33**, W306 - W310

7 Azriel, R. and Gazit, E. (2001) Analysis of the minimal amyloid-forming fragment of the islet amyloid polypeptide - An experimental support for the key role of phenylalanine residue in amyloid formation. Journal of Biological Chemistry **276**, 34156 - 34161

8 DuBay, K. F., Pawar, A. P., Chiti, F., Zurdo, J., Dobson, C. M. and Vendruscolo, M. (2004) Prediction of absolute aggregation rates of amyloidogenic polypeptide chains. Journal of Molecular Biology **341**, 1317 - 1326

9 Dobson, R. C. J., Griffin, M. D. W., Jameson, G. B. and Gerrard, J. A. (2005) The crystal structures of native and (*S*)-lysine-bound dihydrodipicolinate synthase from *Escherichia coli* with improved resolution show new features of biological significance. Acta Crystallographica Section D, Biological Crystallography **D61**, 1116 - 1124

10 DeLano, W. L. (2002) The PyMOL molecular graphics system. DeLano Scientific, San Carlos

11 Villaverde, A. and Mar Carrio, M. (2003) Protein aggregation in recombinant bacteria: biological role of inclusion bodies. Biotechnology Letters **25**, 1385 - 1395

12 Lindner, A. B., Madden, R., Demarez, A., Stewart, E. J. and Taddei, F. (2008) Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation. Proceedings of the National Academy of Sciences USA **105**, 3076 - 3081

13 Dobson, R. C. J., Griffin, M. D. W., Roberts, S. J. and Gerrard, J. A. (2004) Dihydrodipicolinate synthase (DHDPS) from *Escherichia coli* displays partial mixed inhibition with respect to its first substrate, pyruvate. Biochimie **86**, 311 - 315

14 Dobson, R. C. J. (2006) DHDPS stability is reduced at -20 ºC over time. pp. Personal communication, Christchurch

15 Pierce (2005) Technical resource: Protein stability and storage.

16 Karsten, W. E. (1997) Dihydrodipicolinate synthase for *Escherichia coli*: pH dependent changes in the kinetic mechanism and kinetic mechanism of allosteric inhibition by *L*-lysine. Biochemistry **36**, 1730 - 1739

17 Griffin, M. D., Dobson, R. C. J., Pearce, F. G., Antonio, L., Whitten, A. E., Liew, C. K., Mackay, J. P., Trewhella, J., Jameson, G. B., Perugini, M. A. and Gerrard, J. A. (2008) Evolution of quaternary structure in a homotetrameric enzyme. Journal of Molecular Biology **380**, 691 - 703

18 Mirwaldt, C., Korndorfer, I. and Huber, R. (1995) The crystal structure of dihydrodipicolinate synthase from *Escherichia coli* at 2.5 Å resolution. Journal of Molecular Biology **246**, 227 - 239

19 Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallographica Section D: Biological Crystallography **60**, 2126 - 2132

20 Andrews, P. (1964) Estimation of the molecular weights of proteins by sephadex gel filtration. Biochemical Journal **91**, 222 - 233

21 Andrews, P. (1965) The gel filtration behaviour of proteins related to their molecular weights over a wide range. Biochemical Journal **96**, 595 - 606

22 Schuck, P., Perugini, M. A., Gonzales, N. R., Howlett, G. J. and Schubert, D. (2002) Size-distribution analysis of proteins by analytical ultracentrifugation and Lamm equation modeling. Biophysical Journal **78**, 1606 - 1619

23 Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. Biophysical Journal **78**, 1606 - 1619

24 Griffin, M. D. W. (2005) PhD Thesis. Why is DHDPS a tetramer?, University of Canterbury, Christchurch

25 Perugini, M. A., Griffin, M. D. W., Smith, B. J., Webb, L. E., Davin, A. J., Handman, E. and Gerrard, J. A. (2005) Insight into the self-association of key enzymes from pathogenic species. European Biophysical Journal **34**, 469 - 476

26 Pearce, F. G. (2008) Crystal structure of cleaved pET 151/D-TOPO polyhistidine tagged *E. coli* DHDPS. Personal communication, Christchurch

27 Vedadi, M., Niesen, F. H., Allali-Hassani, A., Fedorov, O. Y., Finerty, P. J. J., Wasney, G. A., Yeung, R., Arrowsmith, C., Ball, L. J., Berglund, H., Hui, R., Marsden, B. D., Norlund, P., Sundstrom, M., Weigelt, J. and Edwards, A. M. (2006) Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. Proceedings of the National Academy of Sciences USA **103**, 15835 - 15840

28 Arakawa, T. and Timasheff, S. N. (1984) Mechanism of protein salting in and out by divalent cation salts: Balance between hydration and salt binding. Biochemistry **23**, 5912 - 5923

29 Chi, E. Y., Krishnan, S., Randolph, T. W. and Carpenter, J. W. (2003) Physical stability of proteins in aqueous solution: Mechanism and driving forces in nonnative protein aggregation. Pharmaceutical Research **20**, 1325 - 1336

30 Epps, D. E., Sarver, R. W., Rogers, J. M., Herberg, J. T. and Tomich, P. K. (2001) The ligand affinity of proteins measured by isothermal denaturation kinetics. Analytical Biochemistry **292**, 40 - 50

31 Dümmler, A., Lawrence, A.-M. and de Marco, A. (2005) Simplified screening for the detection of soluble fusion constructs expressed in *E. coli* using a modular set of vectors. Microbial Cell Factories **13**, 34

32 Niesen, F. H., Berglund, H. and Vedadi, M. (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. Nature Protocols **2**, 2212 - 2221

33 Sabate, R., Gallardo, M. and Estelrich, J. (2003) An autocatalytic reaction as a model for the kinetics of the aggregation of β-amyloid. Biopolymers (Peptide Science) **71**, 190 - 195

34 Khurana, R., Coleman, C., Ionescu-Zanetti, C., Carter, S. A., Khrishna, V., Grover, R. K., Roy, R. and Singh, S. (2005) Mechanism of thioflavin T binding to amyloid fibrils. Journal of Structural Biology **151**, 229 - 238

35 Groenning, M., Norrman, M., Flink, J. M., van der Weert, M., Butrinsky, J. T., Schluckebier, G. and Frokjaer, S. (2007) Binding of thioflavin T in insulin amyloid fibrils. Journal of Structural Biology **159**, 483 - 497

36 Saeed, S. M. and Fine, G. (1967) Thioflavin-T for amyloid detection. American Journal of Clinical Pathology **57**, 588 - 593

37 LeVine III, H. (1993) Thioflavine T interaction with synthetic Alzheimer's disease β-amyloid peptides: Detection of amyloid aggregation in solution. Protein Science **2**, 404 - 410

38 Krebs, M. R. H., Bromley, E. H. C. and Donald, A. M. (2005) The binding of thioflavin T to amyloid fibrils: localization and implications. Journal of Structural Biology **149**, 30 - 37

CHAPTER 5

CHARACTERISING AGGREGATION OF *ESCHERICHIA COLI* DHDPS VARIANTS WITH DISRUPTED QUATERNARY STRUCTURE

5.1 INTRODUCTION

The quaternary structure of DHDPS has been shown to be critical for function [1, 2]. However, the role of the quaternary structure on the stability of the protein had not been explored in detail. It was hypothesised that the oligomerisation of the DHDPS monomer might prevent amorphous and β-sheet-specific aggregation. This idea was tested in this chapter.

Previous studies investigating the effect of disruption of the quaternary structure on activity have provided evidence that single amino acid changes at the loose or tight dimer interfaces alter the tetrameric nature of DHDPS (see figure 5.1) [1, 2]. Changing the leucine at position 197 to tyrosine or aspartic acid resulted in a dimeric DHDPS [1]. Altering the tyrosine at position 107 to a tryptophan resulted in a DHDPS that existed in a monomer/dimer/tetramer equilibrium [2]. These variant DHDPS forms had been well studied with regard to their catalytic activity and quaternary structure. However, their propensity to form amorphous or β-sheet-specific aggregate had not been systematically investigated. Anecdotal evidence suggested that these proteins were more prone to aggregation than the wild-type enzyme. The experiments designed to investigate the biophysical properties of these variants, including amorphous and β-sheet-specific aggregation, are described in the following sections.

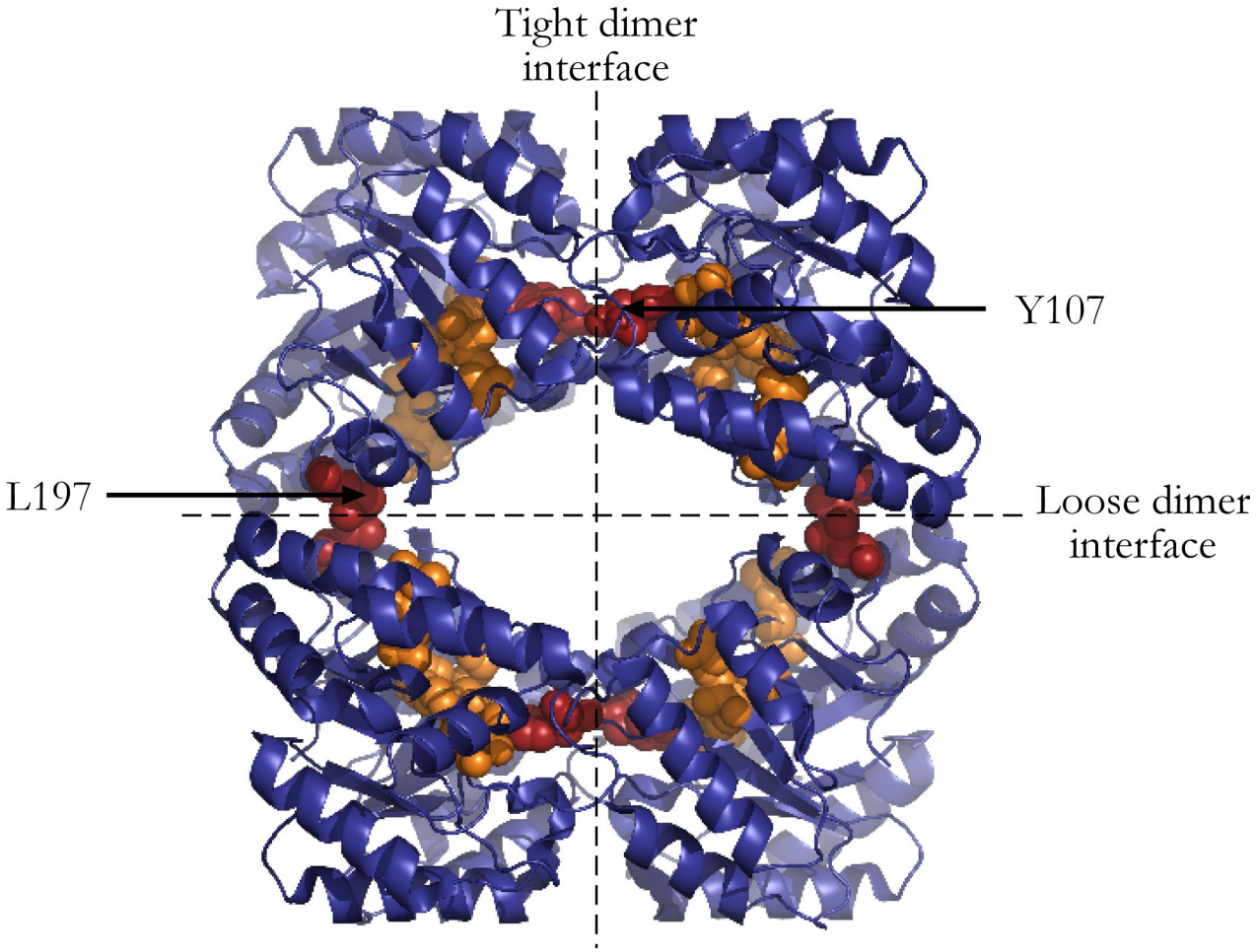


Figure 5.1 –The quaternary structure of *E. coli* DHDPS showing the residues involved in the active site (orange) and the dimer-dimer interfaces. The positions of Y107 and L197 are indicated (red). This image was created using PyMol [3] from the *E. coli* DHDPS PDB file (accession number 1yxc) [4].

Disruption of the quaternary structure of DHDPS exposed the interfaces between the subunits. The residues involved in the interaction of the monomers, some of which are hydrophobic, are shown in figure 5.2. These exposed hydrophobic residues are likely to increase the aggregation propensity of the protein as a whole.

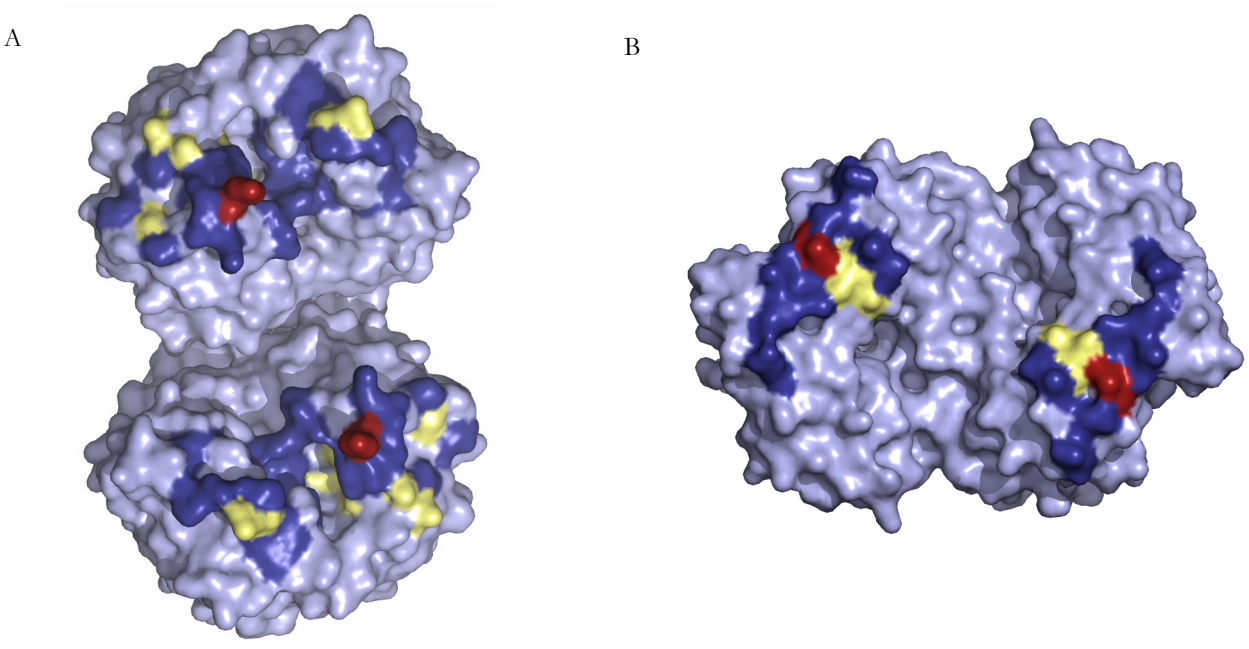


Figure 5.2 –The dimer-dimer interfaces of the *E. coli* DHDPS tetramer showing the residues (dark blue) involved in **(A)** the tight dimer interface showing the position of Y107 (red) that, upon substitution to a tryptophan resulted in DHDPS forming a monomer/tetramer equilibrium [2] and **(B)** the loose dimer interface showing the position of L197 (red), the substitution of which disrupted the quaternary structure of DHDPS, resulting in dimeric species [1]. The yellow residues are the hydrophobic amino acids exposed upon disruption of the quaternary structure. This image was created using PyMol [3] from the *E. coli* DHDPS PDB file (accession number 1yxc) [4].

In addition to characterising the biophysical and aggregation properties of DHDPS Y107W, the combinatorial effects of the pET M11 and pET 151/D-TOPO polyhistidine tags, and residues remaining following cleavage of the pET 151/D-TOPO polyhistidine tag were also tested. As Y107 is located in the active site and involved in catalysis, a second mutation was characterised (Y107F). This amino acid substitution has previously been described [5] and because it did not alter the quaternary structure, it was used as a control. Through comparison of DHDPS Y107F with Y107W, it was possible to establish whether the altered biophysical properties were due to the amino acid substitution or the aberrant quaternary structure caused by the presence of the tryptophan at position 107. Due to the difficulties in purifying DHDPS L197Y (described in section 5.2.1), the characterisation of this variant was limited to its polyhistidine-tagged variants.

The aggregation propensities of the proteins described in this chapter were tested using the Zyggregator algorithm. This predicted that the amino acid substitutions disrupting the quaternary structure DHDPS would have little effect on the aggregation propensity of DHDPS (see appendix 4, section A4.1). However, Zyggregator is limited in the case of amino acid substitutions that result in protein subunits dissociating. The exposure of buried hydrophobic patches is not considered in the calculations, which assume that the proteins are unfolded. Thus the actual aggregation propensity of proteins with aberrant quaternary structures may be higher than predicted. This fact is also a limitation in the interpretation of the data purely in terms of perturbations in quaternary structure since, upon disruption of the interface, hydrophobic patches are exposed which may increase aggregation propensity.

5.2 RESULTS

5.2.1 PURIFICATION

The purifications of DHDPS Y107F, Y107W, its polyhistidine-tagged variants and the polyhistidine-tagged DHDPS L197Y variants were carried out as described in chapter 8, sections 8.4.1, 8.4.4 – 8.4.8. Following purification by ion exchange chromatography (IEX) and hydrophobic interaction chromatography (HIC) the average yield for DHDPS Y107W was 48.6 mg (9.17 units with specific activity of 0.19 units/mg) per litre of bacterial culture. The yield for wild-type DHDPS averaged 19.6 mg (76 units with a specific activity of 3.83 units/mg) per litre of bacterial culture. The dramatic decrease in specific activity is due to the attenuated catalytic ability of DHDPS Y107W [2] which will be discussed in section 5.2.2.1. The addition of the polyhistidine tags reduced the yield compared to DHDPS Y107W and the specific activity compared to the wild-type enzyme. pET M11 DHDPS Y107W yielded 23.7 mg (3.3 units with a specific activity of 0.14 units/mg) per litre of bacterial culture. pET 151/D-TOPO DHDPS Y107W yielded 40.3 mg (2.4 units with a specific activity of 0.06 units/mg) per litre of bacterial culture. Cleavage of the pET 151/D-TOPO polyhistidine tag from DHDPS Y107W was carried out according to the protocol described in chapter 8, section 8.4.9 and yielded 66 % of the original protein present (26.6 mg/L, equivalent to 3.1 units with a specific activity of 0.12 units/mg). The cause of this loss is discussed in chapter 3, section 3.3.2. DHDPS Y107F yielded an average 33.1 mg (0.48 units with specific activity of 0.01 units/mg) per litre of bacterial culture. SDS-PAGE revealed that all the variants were the major protein in solution and in

most cases a single species as judged by SDS-PAGE with Coomassie blue staining (see figure 5.3). Thus catalytic attenuation due to the amino acid substitutions at position Y107 is the likely cause of the reduced specific activities.

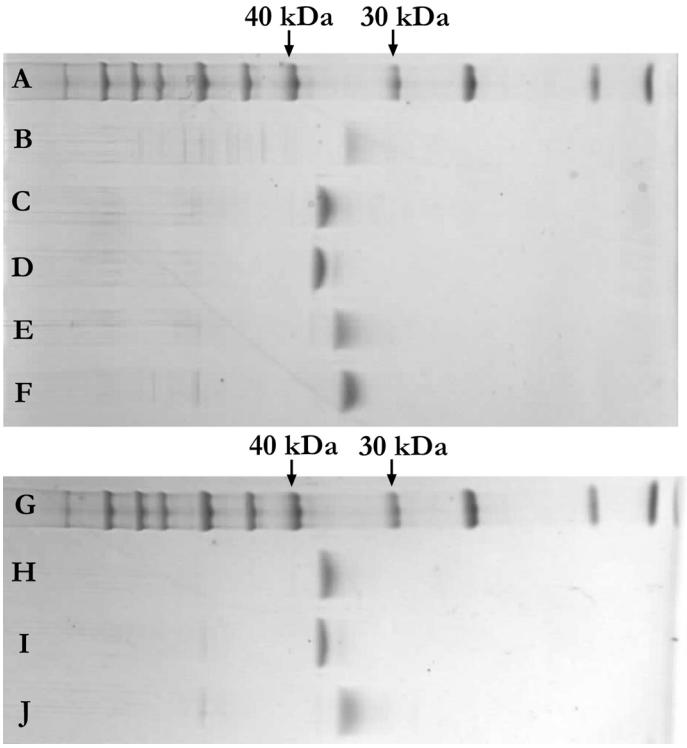


Figure 5.3 – SDS-PAGE. The lanes were loaded with **(A)** molecular weight marker with relevant bands indicated, **(B)** DHDPS Y107W, **(C)** pET M11 DHDPS Y107W, **(D)** pET 151-D-TOPO DHDPS Y107W, **(E)** cleaved pET 151-D-TOPO DHDPS Y107W, **(F)** DHDPS Y107F, **(G)** molecular weight marker with relevant bands indicated, **(H)** pET M11 DHDPS L197Y, **(I)** pET 151-D-TOPO DHDPS L197Y and **(J)** cleaved pET 151-D-TOPO DHDPS L197Y.

The amino acid substitution at position 197 dramatically reduced yields. The yield for DHDPS L197Y was 1.2 mg (0.04 units with specific activity of 0.04 units/mg) per litre of bacterial culture. For this reason, and the anecdotal evidence that DHDPS L197Y was unstable [6], it was not used in further experiments. However, upon addition of the polyhistidine tags to DHDPS L197Y, sufficient protein was obtained to carry out further assays. The yield for pET M11 DHDPS L197Y averaged 19.8 mg (0.73 units with specific activity of 0.04 units/mg) per litre of culture. The yield for pET 151/D-TOPO DHDPS L197Y averaged 22.6 mg (3.2 units with specific activity of 0.14 units/mg) per litre of culture. The cleavage of the pET 151/D-TOPO polyhistidine tag yielded 57 % of the original protein, averaging 12.9 mg (0.78 units with a specific activity of 0.06 units/mg) per litre of

bacterial culture. As discussed below, the reduced specific activity is most likely due to impaired catalytic ability of the enzyme, a postulate supported by the results of SDS PAGE (see figure 5.3) which showed few contaminants in the purified protein samples.

5.2.2 BIOPHYSICAL CHARACTERISATION

Using the techniques described in chapter 2, DHDPS Y107W, L197Y, their polyhistidine-tagged variants and DHDPS Y107F were characterised. The outcomes of these assays are described in the following sections and the results are compared to those obtained for wild-type DHDPS.

5.2.2.1 KINETICS

DHDPS variants, containing amino acid substitutions at positions 107 and 197, exhibit very different kinetic parameters to the wild-type enzyme. The k_{cat}^{app} values of Y107W and its polyhistidine variants are significantly reduced (see table 5.1). The results of the pseudo-single substrate kinetic analysis carried out during the course of this thesis are similar to previously published values for DHDPS Y107W (k_{cat} 6.6 ± 0.4 s^{-1}, $K_{m\ Pyr}$ 1.5 ± 0.2 mM and $K_{m\ (S)\text{-ASA}}$ 0.39 ± 0.04 mM [2]). The attenuation of catalytic functionality of Y107W could be attributed to both a loss of conformational rigidity leading to weaker binding of the substrates [1] and the involvement of the tyrosine at position 107 in the active site [2, 5]. The replacement of this tyrosine with a phenylalanine also decreased the k_{cat}^{app} significantly [5] with the results of the assays consistent with published data (k_{cat} 10.8 ± 0.2 s^{-1}, $K_{m\ Pyr}$ 0.16 ± 0.01 mM and $K_{m\ (S)\text{-ASA}}$ 0.58 ± 0.03 mM [5]). The results of the assay performed during the course of this thesis corroborated previously published data [2, 5] and confirmed that Y107 is a key residue in catalysis [5]. The presence of the extra amino acids associated with the polyhistidine tags, in combination with the amino acid substitution at position 107, reduced the k_{cat}^{app} further; a result analogous to those described in chapter 3, section 3.3.3.1 (although the precise values for the k_{cat}^{app} are difficult to establish where the K_m^{app} is high). The Michaelis constants for both pyruvate and (*S*)-ASA are increased for DHDPS Y107W. The presence of the polyhistidine tags increased the K_m^{app} with respect to pyruvate further whilst decreasing the K_m^{app} for (*S*)-ASA.

The kinetic parameters of the polyhistidine-tagged variants of L197Y are also altered compared to the non-tagged DHDPS L197Y. Griffin *et al.* [1] showed that the k_{cat} for DHDPS L197Y to be 1.11

± 0.05 s^{-1} (1.4 % of the k_{cat} for wild-type) and the Michaelis constants for pyruvate and (*S*)-ASA were 1.11 ± 0.10 mM and 0.18 ± 0.01 mM respectively. In addition to this, the kinetic mechanism of DHDPS L197Y was found to change from the ping-pong model observed for wild type to a ternary complex mechanism [1]. The increase in K_m values for pyruvate and reduced activity were attributed to a trapped substrate analogue at the active site [1]. The polyhistidine tags on DHDPS L197Y generally caused little change in the k_{cat}^{app} values compared to the published values for non-tagged L197Y; however, the pET 151/D-TOPO did increase the k_{cat}^{app} and the K_m^{app} with respect to pyruvate although the activity was still severely attenuated compared to the wild-type enzyme. The Michaelis constants for most variants were similar to wild-type. The exceptions to this were the increase in K_m^{app} with respect to pyruvate for the pET 151/D-TOPO DHDPS L197Y and its cleaved variant and the reduction in K_m^{app} with respect to pyruvate for pET M11 DHDPS L197Y compared to the published results for the non-tagged variant.

	k_{cat}^{app} with respect to Pyruvate (s^{-1})	**$K_{m\ Pyr}^{app}$ (mM)**	**k_{cat}^{app} with respect to (*S*)-ASA (s^{-1})**	**$K_{m\ (S)\text{-}ASA}^{app}$ (mM)**
Wild-type DHDPS	118.1 ± 2.8	0.3 ± 0.0	123.2 ± 4.8	0.1 ± 0.0
DHDPS Y107W	6.04 ± 1.39	3.69 ± 0.35	5.81 ± 0.43	1.22 ± 0.09
pET M11 DHDPS Y107W	0.68 ± 0.07	12.45 ± 0.33	8.14 ± 0.66	0.49 ± 0.11
pET 151/D-TOPO DHDPS Y107W	1.11 ± 0.10	6.43 ± 0.15	2.65 ± 0.18	0.61 ± 0.11
Cleaved pET 151/D-TOPO DHDPS Y107W	1.43 ± 0.21	12.19 ± 0.47	5.85 ± 0.28	0.37 ± 0.05
DHDPS Y107F	0.27 ± 0.09	2.44 ± 0.22	0.63 ± 0.13	0.50 ± 0.07
pET M11 DHDPS L197Y	1.81 ± 0.39	0.15 ± 0.01	0.49 ± 0.03	0.31 ± 0.05
pET 151/D-TOPO DHDPS L197Y	5.55 ± 0.69	5.54 ± 0.26	3.06 ± 0.15	0.36 ± 0.05
Cleaved pET 151/D-TOPO DHDPS L197Y	0.87 ± 0.19	2.28 ± 0.12	2.92 ± 0.16	0.53 ± 0.08

Table 5.1 – The kinetic parameters for DHDPS Y107W, L197Y, their polyhistidine-tagged variants and Y107F. Errors represent the SEM of at least three replicates.

Existing literature attributes the reduction of catalytic efficiency of L197Y to the increased flexibility of the dimeric enzyme [1]. It was postulated that the tetrameric nature of DHDPS was an evolutionarily driven adaptation to minimise the movement between the monomers, thus improving catalytic efficiency [1]. This hypothesis was put into question by the discovery of an active, dimeric DHDPS in *S. aureus* and MRSA as discussed in chapter 1, section 1.10. In the following sections, the alternative hypothesis (that the tetrameric nature of DHDPS protects it from aggregation), is explored.

5.2.2.2 CRYSTALLOGRAPHIC ANALYSIS

The crystal structure of DHDPS Y107W [2] (PDB accession number 3den) and Y107F [5] (PDB accession number 1s5v) demonstrate no change in tertiary or quaternary structure compared to wild-type DHDPS (see figures 5.4 and 5.5). The structure of L197Y was solved by Griffin *et al.* [1] (PDB accession number 2ojp) and showed that, while the dimeric architecture was the same, the tetrameric nature of the protein was disrupted, resulting in the formation of a discrete tight-dimer (see figures 5.6 and 5.7).

The formation of crystals of the polyhistidine-tagged variants of DHDPS Y107W and L197Y was not successful either under standard conditions (see chapter 2, section 2.3.2.2) or under any conditions in the sparse-matrix screens (Hampton Research screens HR2-110 and HR2-112). This was consistent with the unsuccessful attempts to crystallise the polyhistidine-tagged variants of wild-type DHDPS.

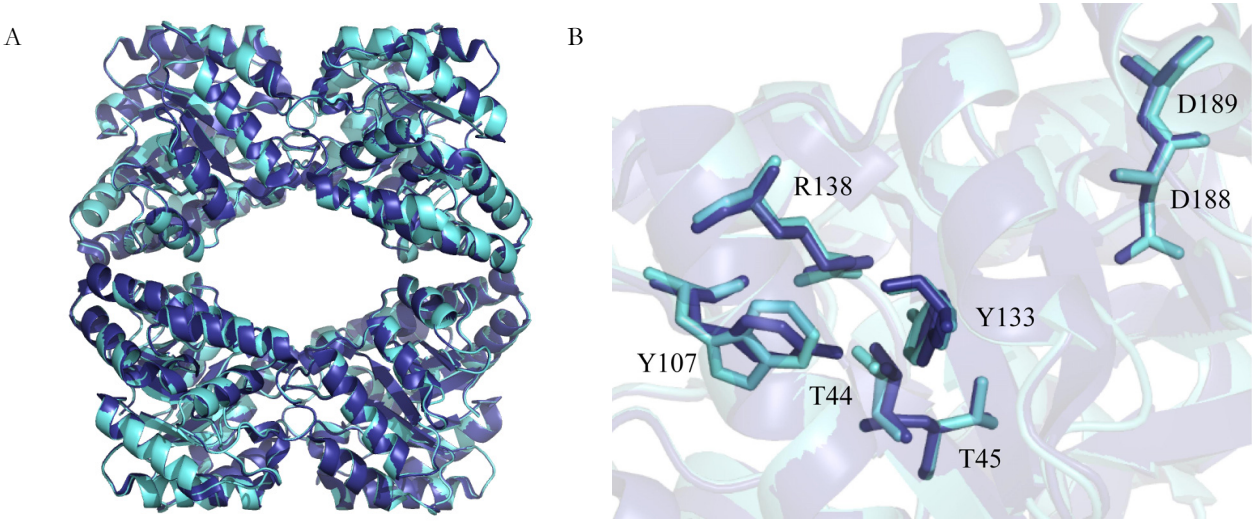


Figure 5.4 – **(A)** Alignment of the structure of wild-type DHDPS (dark blue) (PDB accession number 1yxc) [4] with the structure of DHDPS Y107W (light blue) (PDB accession number 3den) [2] exhibited no significant alterations to either the tertiary or quaternary structure (rmsd = 0.226 Å). **(B)** Superposition of the residues involved in the active site demonstrated no significant structural changes in the crystalline state and confirms the amino acid substitution at position 107. These images were created using PyMol [3].

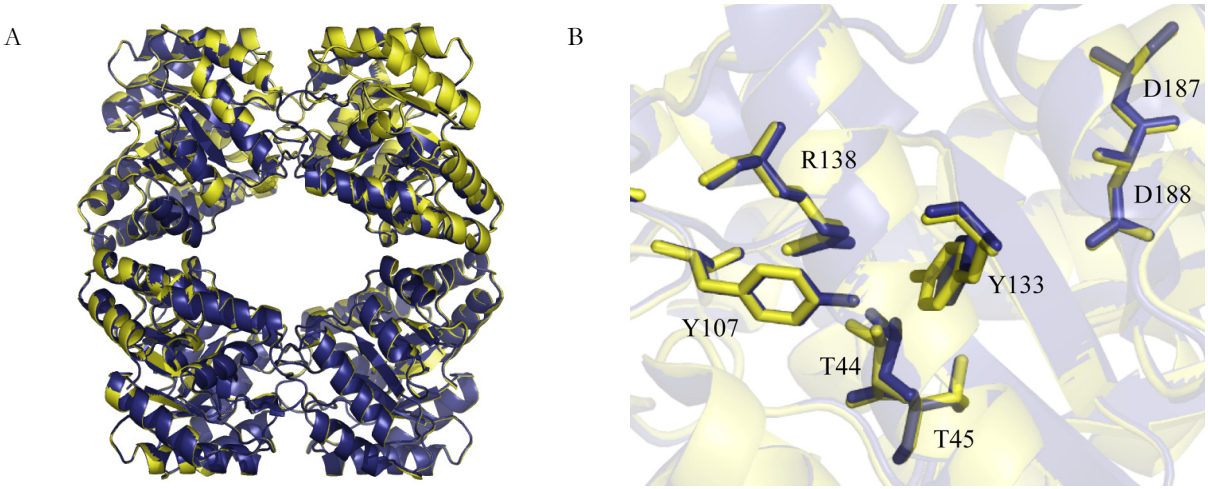


Figure 5.5 – **(A)** Alignment of the structure of wild-type DHDPS (blue) (PDB accession number 1yxc) [4] with the structure of DHDPS Y107F (yellow) (PDB accession number 1s5v) [5] exhibited no significant alterations to either the tertiary or quaternary structure (rmsd = 0.192 Å). **(B)** Superposition of the residues involved in the active site demonstrated no significant structural changes in the crystalline state and confirms the amino acid substitution at position 107. These images were created using PyMol [3].

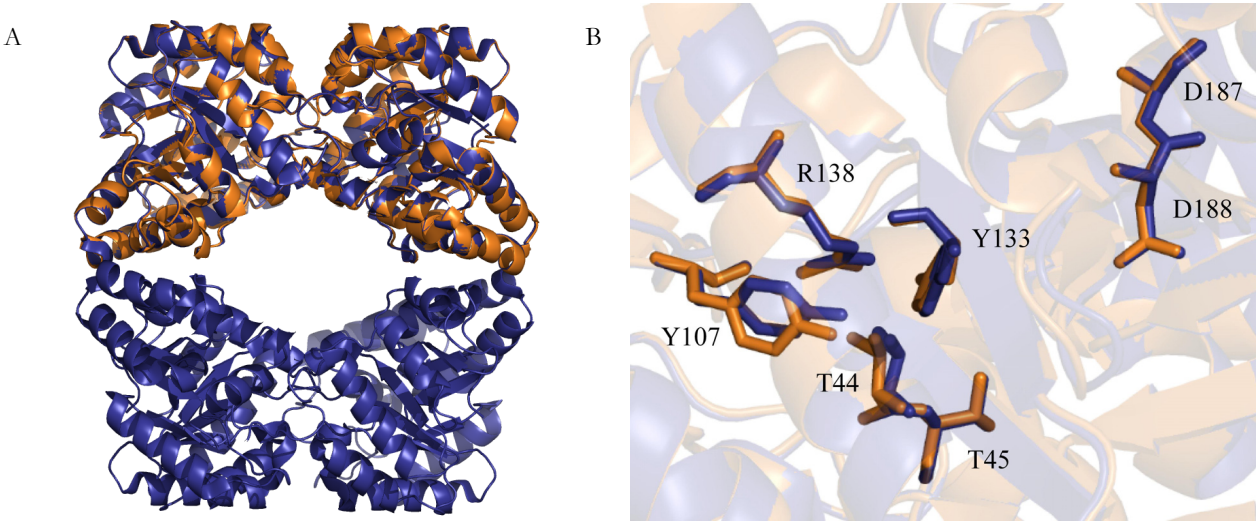


Figure 5.6 – **(A)** Alignment of the structure of wild-type DHDPS (blue) (PDB accession number 1yxc) [4] with the structure of DHDPS L197Y dimer (orange) (PDB accession number 2ojp) [5] exhibited no significant alterations to either the tertiary or quaternary structure (rmsd = 0.217 Å). **(B)** Superposition of the residues involved in the active site demonstrated no significant structural changes in the crystalline state. These images were created using PyMol [3].

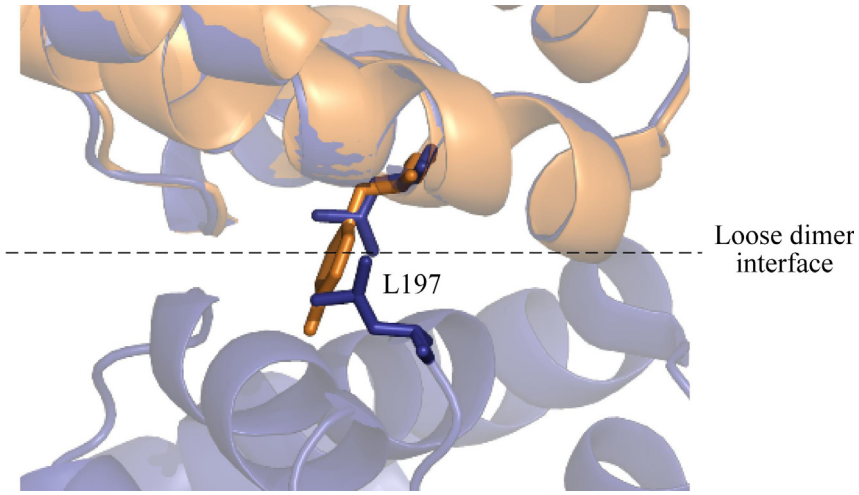


Figure 5.7 –The superposition of the structures confirms the amino acid substitution at position 197. This image were created using PyMol [3] and structures of wild-type DHDPS (blue) (PDB accession number 1yxc) [4] and DHDPS L197Y dimer (orange) (PDB accession number 2ojp) [5].

The crystal structures show no significant change in the tertiary structure of DHDPS upon mutation of the residues at positions 107 and 197 to a tryptophan and a tyrosine or phenylalanine, respectively. The effects of the amino acid substitutions on quaternary structure are well documented [1, 2] and are discussed below.

5.2.2.3 ANALYTICAL GEL PERMEATION LIQUID CHROMOTOGRAPHY

The molecular masses of the variants were calculated from their V_e and the calibration plot described in chapter 2, section 2.3.2.3. The results confirm that all the Y107W variants have disrupted quaternary structures, with monomeric molecular masses (see figure 5.8 and table 5.2). Pearce *et al.* [2] found DHDPS Y107W was in a monomer/dimer/tetramer equilibrium with the tetramer being the predominant species in 20 mM Tris (pH 8), 150 mM NaCl. They found that the dimer was difficult to resolve due to its low, slow exchange constants [2]. The gel filtration of DHDPS Y107W carried out during the course of this research indicated that in 20 mM Tris (pH 7) and 100 mM NaCl, the predominant species was monomeric with a small population of larger structures explaining the presence of only two peaks by gel filtration. As the buffer conditions for the published data were different (pH 8, 150 mM NaCl), the disparity between the published results and those

described in this thesis suggests that the buffer conditions influence the monomer/tetramer equilibrium.

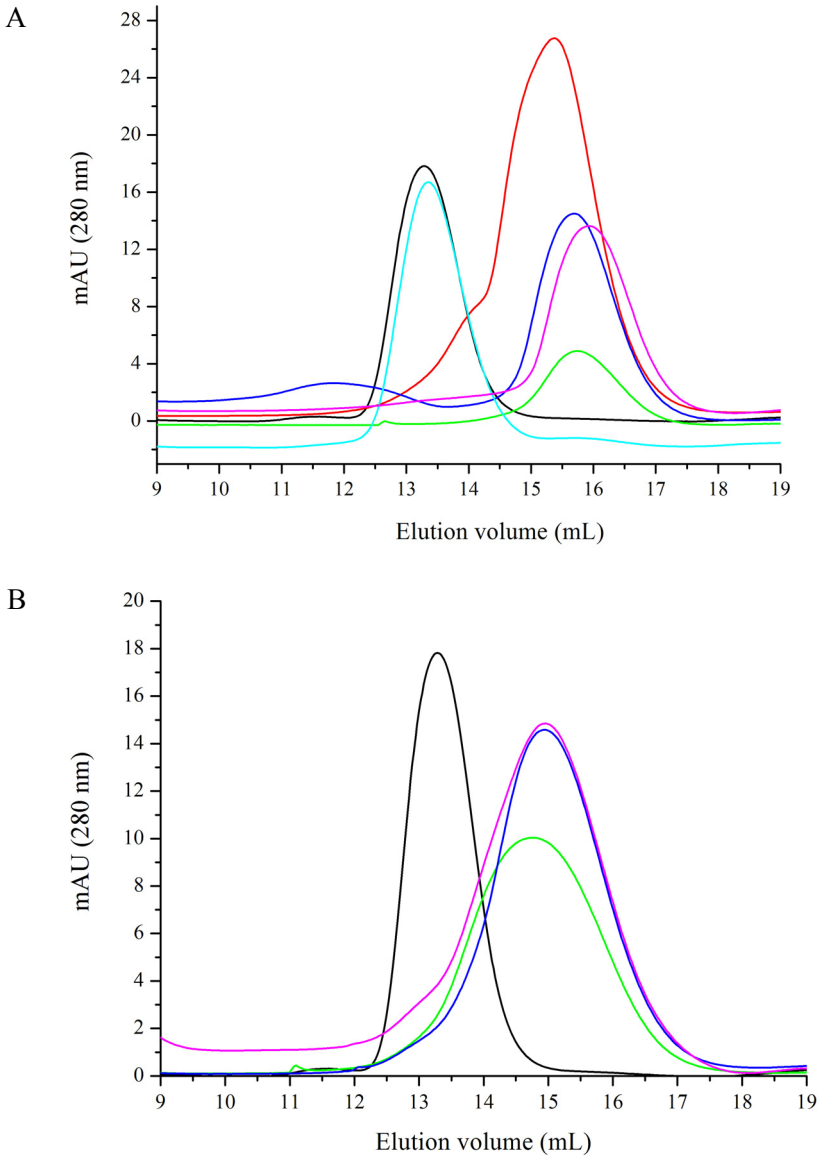


Figure 5.8 – Elution profiles for **(A)** wild-type DHDPS (black), DHDPS Y107W (red) pET M11 DHDPS Y107W (green), pET151/D-TOPO DHDPS Y107W (dark blue), cleaved pET151/D-TOPO DHDPS Y107W (pink) and DHDPS Y107F (light blue). **(B)** wild-type DHDPS (black), pET M11 DHDPS L197Y (green), pET151/D-TOPO DHDPS L197Y (blue) and cleaved pET151/D-TOPO DHDPS L197Y (pink) (at 0.5 mg/mL in 20 mM Tris (pH 7), 100 mM NaCl). The chromatograms were produced by plotting the absorbance at 280 nm as a function of elution volume (V_e). Like wild-type DHDPS, Y107F exhibited an elution volume of 13.3 mL indicating that it is tetrameric. The other variants exist as either monomers or dimers in solution. The calculated molecular masses are shown in table 5.2.

Variant	Calculated molecular mass (kDa)
Wild-type DHDPS	126.7
DHDPS Y107W	34.5
pET M11 DHDPS Y107W	35.0
pET151/D-TOPO DHDPS Y107W	31.0
cleaved pET151/D-TOPO DHDPS Y107W	32.0
DHDPS Y107F	129.6
pET M11 DHDPS L197Y	57.7
pET151/D-TOPO DHDPS L197Y	52.7
cleaved pET151/D-TOPO DHDPS L197Y	51.9

Table 5.2 – The molecular masses of wild-type DHDPS and its variants as calculated from the elution volumes obtained from the curves shown in figure 5.7 and the calibration plot shown in chapter 2, section 2.3.2.3.

The presence of the polyhistidine tags and the residues remaining following cleavage of the pET 151/D-TOPO polyhistidine tag appear to result in the protein population existing primarily as monomers, with calculated molecular masses similar to that of the tagged wild-type monomers. Cleaved pET 151/D-TOPO DHDPS Y107W eluted at different volumes from the other variants. However, it had a molecular mass closer to monomeric DHDPS. The underlying cause of this may have been due to the behaviour of the protein on the column. As expected, Y107F had an elution volume and thus a molecular mass similar to wild-type DHDPS.

The molecular mass of the polyhistidine-tagged L197Y variants were more similar to the predicted mass of tagged, dimeric DHDPS (57.7 kDa for pET M11 DHDPS, 52.7 kDa for pET 151/D-TOPO DHDPS and 51.9 kDa for cleaved pET 151/D-TOPO DHDPS. These results showed good agreement with those published by Griffin *et al.* [1] who found DHDPS L197Y to be a dimer. The presence of the extra amino acids associated with the polyhistidine tags did not affect the dimeric nature of DHDPS L197Y.

Many of the results are outside the 10 % error margin [7, 8] associated with the calculation of molecular masses from analytical gel permeation liquid chromatography data. This indicates that some of their biophysical properties may be altered. These modifications may be a result of regions of hydrophobicity being exposed upon disruption of the quaternary structure, affecting the protein's behaviour on the column. This is supported by the variable heights of the protein peaks, which suggest that some aggregation was occurring during the course of the assay. The effects of the

variability in quaternary structure on other biophysical parameters are described in following sections.

5.2.2.4 ANALYTICAL ULTRA-CENTRIFUGATION

Sedimentation velocity AUC analyses were performed on pET M11 DHDPS Y107W and L197Y (see figures 5.9 and 5.10 respectively) in order to confirm the result obtained by analytical gel permeation liquid chromatography (section 5.2.2.) as described in previous chapter 2, 2.3.2.3. The sedimenting boundaries for pET M11 DHDPS Y107W spread over time, suggesting the samples were not homogeneous (figure 5.9(B)). The quality of the nonlinear least squares best-fit was affirmed by the low rmsd number of 0.022 and the random residual distribution data (figure 5.9(A)). The $c(s)$ data showed peaks with ordinate maxima corresponding to sedimentation coefficients of 2.8 S and 5.8 S. (figure 5.9(C)). These values cannot be compared to results obtained for DHDPS variants with only one oligomeric state [1, 6, 9] because the sedimentation ($c(s)$) analysis assumes a non-interacting species. Thus, the average frictional ratio would not correspond to either of the species, resulting in inaccurate sedimentation coefficients; however, the values correspond with published values for Y107W (2.9 S and 6.5 S [2]). It can be concluded that there are several oligomeric species of tagged DHDPS Y107W present in solution, a result similar to that observed for the non-tagged species [2]. The low, slow exchange constants of the subunits of DHDPS Y107W described by Pearce *et al.* also explains the observation of only two species by AUC [2]. The differences between the techniques in terms of the ratio between the monomer and tetramer have previously been attributed to the interaction of the proteins with the gel matrix during the analytical gel permeation liquid chromatography. AUC does not support the results of the analytical gel permeation liquid chromatography, which demonstrated that the polyhistidine-tagged variants existed as monomers in solution. This discrepancy in results is most likely due to the behaviour of the proteins in the gel matrix of the column. These results suggest that the polyhistidine tag can affect the quaternary structure of DHDPS Y107W under some conditions.

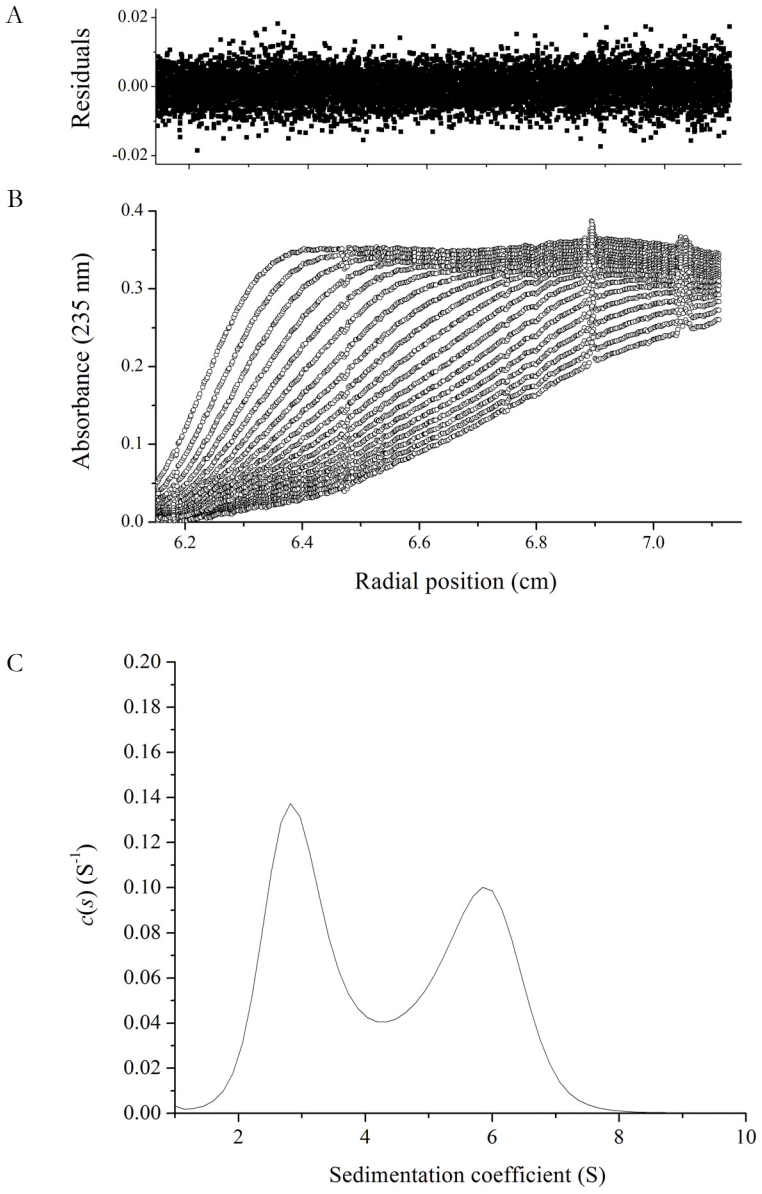


Figure 5.9 – The sedimentation velocity of pET M11 DHDPS Y107W at 0.1 mg/mL in 20 mM Tris (pH 7), 100 mM NaCl at 18 ºC. **(A)** Residuals for the nonlinear least squares best fit are plotted as a function of the radial position from the axis of rotation (cm). **(B)** Absorbance at 235 nm plotted as a function of the radial position from the axis of rotation (cm). Both the raw data (symbols) and fitted data (solid line) are plotted at time intervals of 6 minutes. Raw data are overlaid with the nonlinear least squares best fit to a continuous size-distribution model. **(C)** The continuous sedimentation coefficient ($c(s)$) distribution plotted as a function of the sedimentation coefficient (S). The fit obtained was using a resolution of 100 species between s_{min} of 1 S and s_{max} of 10 S with P = 0.95, $\bar{\nu}$ = 0.7364, ρ = 1.005 g/mL, η = 1.0214 cp and f/f_0 = 1.23.

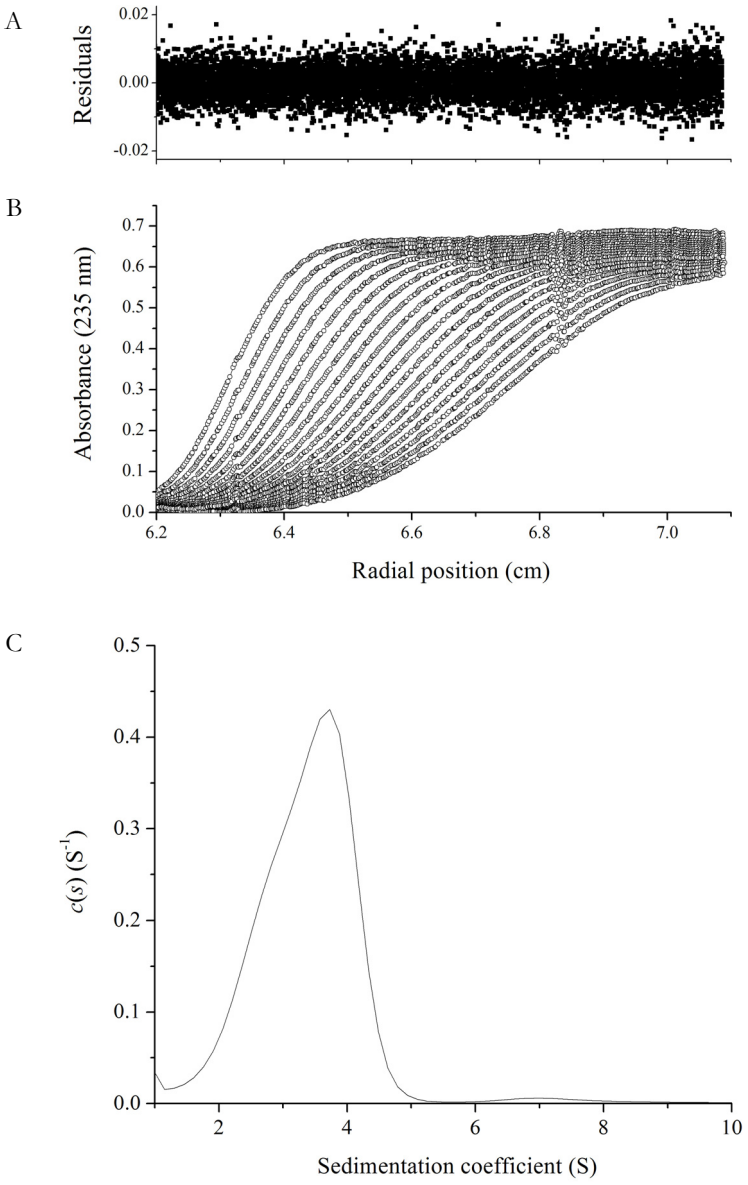


Figure 5.10 – The sedimentation velocity of pET M11 DHDPS L197Y at 0.1 mg/mL in 20 mM Tris (pH 7), 100 mM NaCl at 18 ºC. **(A)** Residuals for the nonlinear least squares best fit are plotted as a function of the radial position from the axis of rotation (cm). **(B)** Absorbance at 235 nm plotted as a function of the radial position from the axis of rotation (cm). Both the raw data (symbols) and fitted data (solid line) are plotted at time intervals of 6 minutes. Raw data are overlaid with the nonlinear least squares best fit to a continuous size-distribution model. **(C)** The continuous sedimentation coefficient ($c(s)$) distribution plotted as a function of the sedimentation coefficient (S). The fit obtained was using a resolution of 100 species between s_{min} of 1 S and s_{max} of 10 S with P = 0.95, $\bar{\nu}$ = 0.7367, ρ = 1.005 g/mL, η = 1.0214 cp and f/f_0 = 1.41.

The sedimenting boundaries for pET M11 DHDPS L197Y did not undergo any significant spreading with respect to time suggesting that the sample was homogeneous (figures 5.10(B)). This was confirmed by fitting the data to a continuous size-distribution model [10, 11]. The $c(s)$ data showed a primary peak with an ordinate maximum corresponding to a sedimentation coefficient of 3.7 S (figures 5.10(C)). This value was consistent with results previously documented for DHDPS L197Y [1, 6] and indicates that pET M11 DHDPS L197Y exists primarily as a dimer. The $c(s)$ data are slightly aberrant from the Gaussian distribution usually observed for DHDPS. This may indicate the presence of a smaller species and could potentially be due to the presence of a small population of monomer in solution. The quality of the nonlinear least squares best-fit was affirmed by the rmsd number (0.109) and the random residual distribution data (figures 5.10(A)).

5.2.2.5 DIFFERENTIAL SCANNING FLUORIMETRY

DSF was used to assess the thermal denaturation temperatures of DHDPS Y107F, Y107W, and L197Y, along with the polyhistidine-tagged variants of the latter. The results obtained were compared to those of wild-type DHDPS (chapter 2, section 2.3.2.5). Statistical analysis was carried out as described in chapter 3, section 3.3.3.4. The analysis of deviance table for all DHDPS variants is presented in appendix 1, table A1.2.

The effects of the amino acid substitutions and the combinatorial effects of the polyhistidine tags on the thermal denaturation temperature of DHDPS in 100 mM phosphate buffer are shown in figures 5.11 – 5.14 with the data presented in their entirety in appendix 4, section A4.2. The results indicated that the thermal stability of DHDPS was reduced through disruption of the quaternary structure, with DHDPS Y107W, L197Y and their polyhistidine-tagged variants exhibiting significantly lower thermal denaturation temperatures than wild-type (figure 5.11 and 5.12). Y107F demonstrated a thermal denaturation temperature similar to wild-type, supporting the hypothesis that it is the disruption of the tetramer that decreases stability. The results of this assay also indicate that the presence of the extra amino acids associated with the polyhistidine tags do not play a significant role in the stability of the proteins, supporting the results described in chapter 3, section 3.3.3.4. The underlying causes of inconsistent values obtained for the maximum RFU are discussed in chapter 3, section 3.3.3.4.

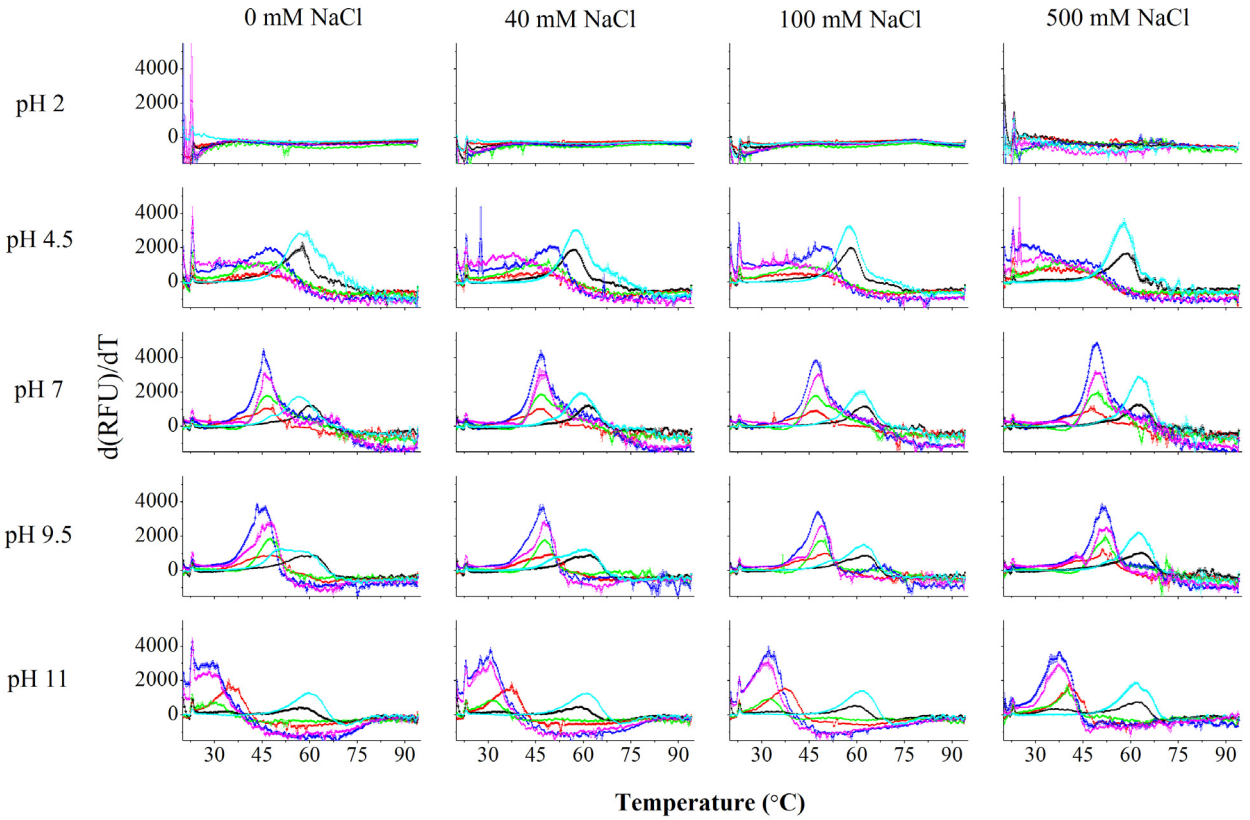


Figure 5.11 – The thermal denaturation profiles of wild-type DHDPS (black), DHDPS Y107W (red), pET M11 DHDPS Y107W (green), pET 151/D-TOPO DHDPS Y107W (dark blue), cleaved pET 151/D-TOPO DHDPS Y107W (pink) and DHDPS Y107F (light blue) (at 0.5 mg/mL in 100 mM phosphate buffer, pH 7, 40 mM NaCl) as monitored by SYPRO orange fluorescence. The data plotted are the derivatives of the increase in fluorescence monitored by the BioRad IQ5 (presented in appendix A4, figure A4.1). The peaks indicate the temperature at which the protein is unfolding the most rapidly and is recorded as the mean thermal denaturation temperature. The data plotted are the mean values ± SEM (n=3).

In keeping with the results from previous chapters, the effect of salt on the thermal denaturation temperature of DHDPS Y107F, Y107W, L197Y and the polyhistidine-tagged variants of the latter was minimal. The effect of pH was significant with all variants exhibiting high initial fluorescence at pH 2, characteristic of all DHDPS variants described thus far. As discussed in previous chapters, this indicated that the protein had exposed hydrophobic patches prior to the heating regime [12]. The Y107W variants had the highest thermal stability at pH 7 and 9.5, and exhibited decreased thermal stability at pH 4.5 and 11 (see figure 5.13). However, some of the Y107W variants underwent thermal denaturation at pH 4.5 at similar temperatures to pH 7 and pH 9.5. The L197Y variants

generally exhibited similar trends in thermal denaturation as described for other DHDPS species; however, at pH 4.5 the analyses of the data for pET M11 and pET 151/D-TOPO DHDPS L197Y were complicated by high initial fluorescence and an indistinct peak in the fluorescence data (see appendix 4, figure A4.2). The variability in thermal denaturation data at pH 4.5 exhibited by all DHDPS species with aberrant quaternary structure may indicate altered properties upon protonation. This was not a result of a change in the pI values of the Y107W and L197Y variants, which were similar to the wild-type protein and its polyhistidine-tagged variants.

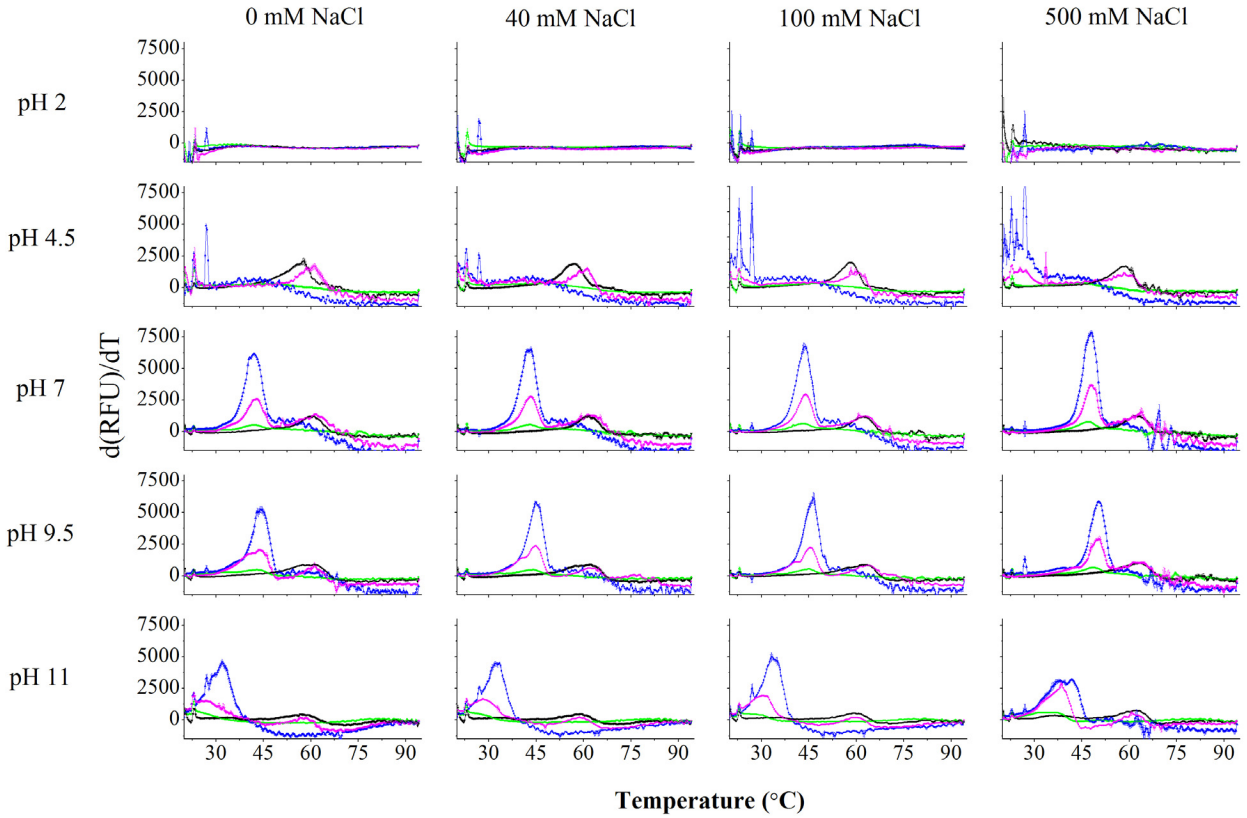


Figure 5.12 – The thermal denaturation profiles of wild-type DHDPS (black), pET M11 DHDPS L197Y (green), pET 151/D-TOPO DHDPS L197Y (dark blue), and cleaved pET 151/D-TOPO DHDPS L197Y (pink) (at 0.5 mg/mL in 100 mM phosphate buffer, pH 7, 40 mM NaCl) as monitored by SYPRO orange fluorescence. The data plotted are the derivatives of the increase in fluorescence monitored by the BioRad IQ5 (presented in appendix A4, figure A4.2). The peaks indicate the temperature at which the protein is unfolding the most rapidly and is recorded as the mean thermal denaturation temperature. The data plotted are the mean values ± SEM (n=3).

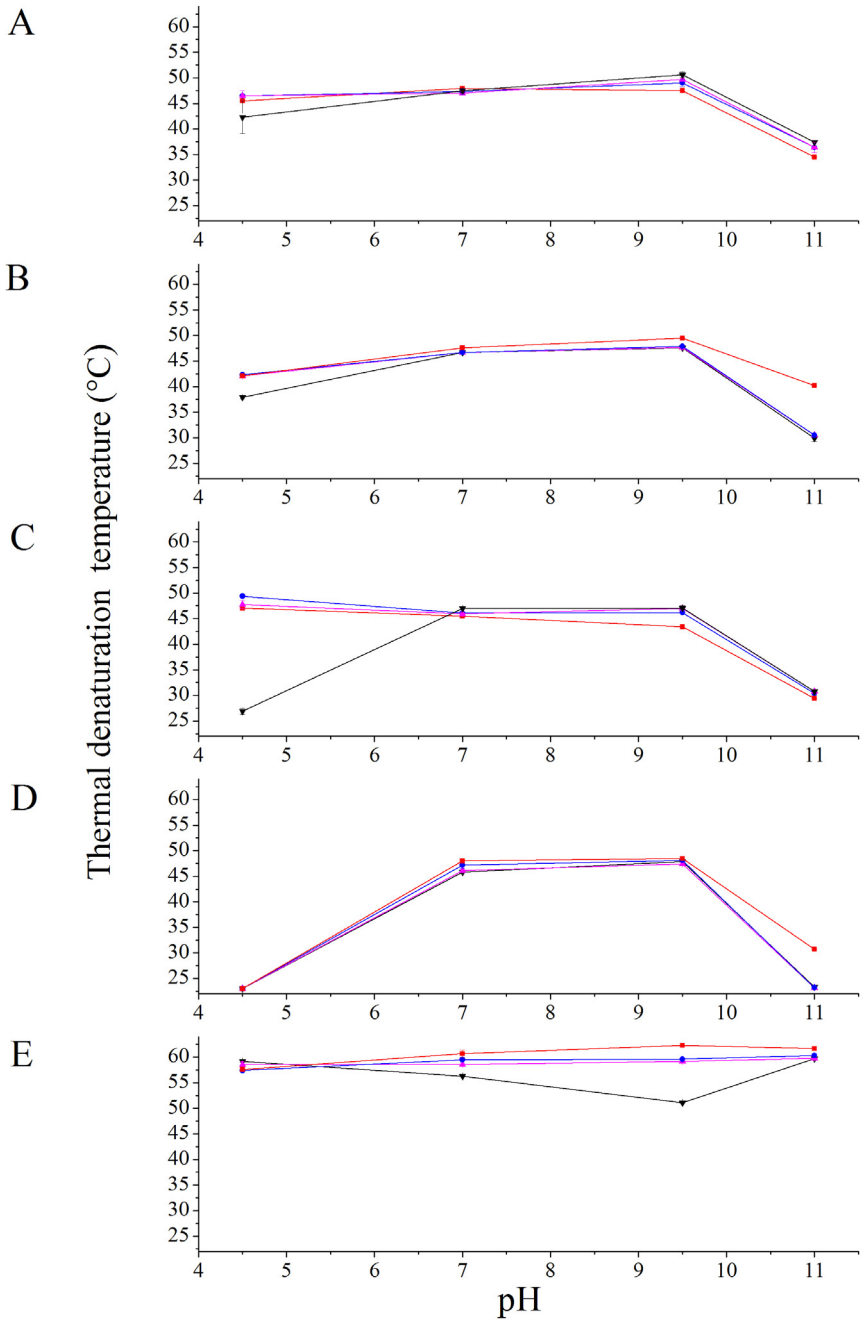


Figure 5.13 – The primary thermal denaturation temperatures of **(A)** DHDPS Y107W, **(B)** pET M11 DHDPS Y107W, **(C)** pET 151/D-TOPO DHDPS Y107W, **(D)** cleaved pET 151/D-TOPO DHDPS Y107W and **(E)** DHDPS Y107F in 100 mM phosphate buffer at 0 mM NaCl (red), 40 mM NaCl (blue), 100 mM NaCl (pink) and 500 mM NaCl (black) plotted as a function of pH. The mean thermal denaturation temperatures plotted were calculated from derivations of the fluorescence data. The error bars represent the SEM.

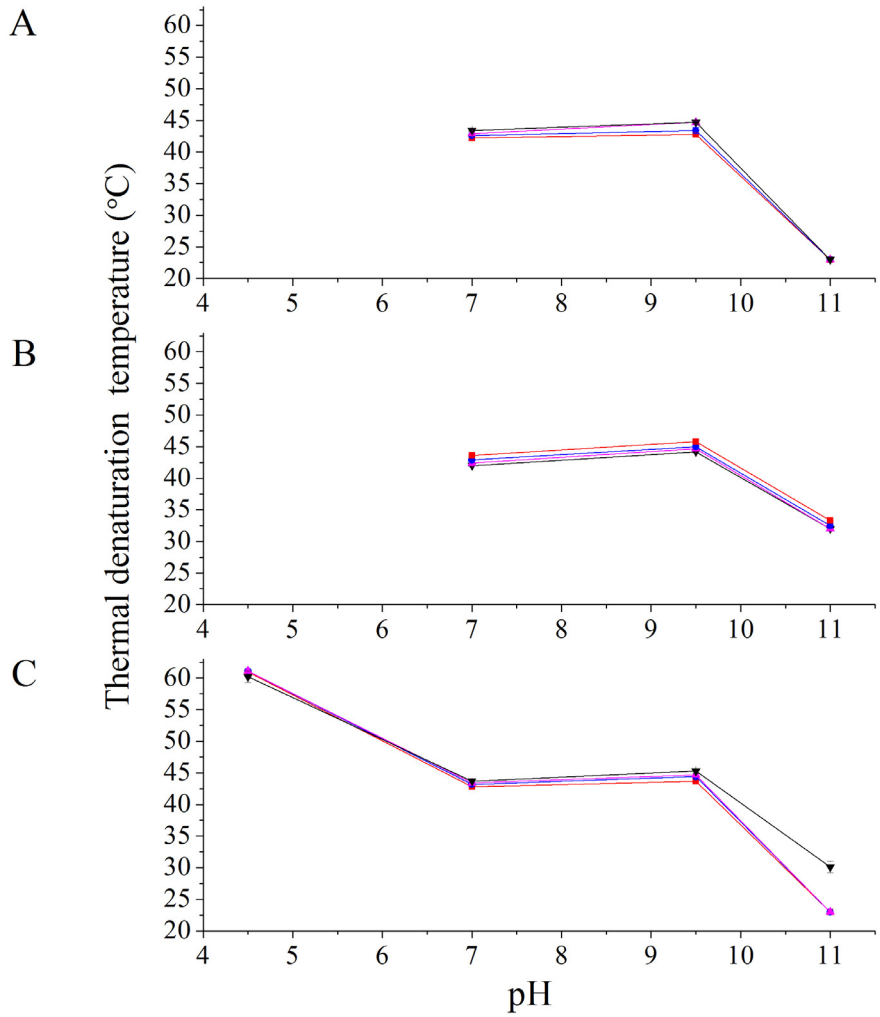


Figure 5.14 – The primary thermal denaturation temperatures of **(A)** pET M11 DHDPS L197Y, **(B)** pET 151/D-TOPO DHDPS L197Y, and **(C)** cleaved pET 151/D-TOPO DHDPS L197Y a in 100 mM phosphate buffer at 0 mM NaCl (red), 40 mM NaCl (blue), 100 mM NaCl (pink) and 500 mM NaCl (black) plotted as a function of pH. The mean thermal denaturation temperatures plotted were calculated from derivations of the fluorescence data. Note that the data for pH 4.5 was not able to be used to calculate the thermal denaturation temperatures of pET M11 and pET 151/D-TOPO DHDPS L197Y. The error bars represent the SEM.

Some of the variants exhibited biphasic thermal denaturation data under some conditions (see appendix 4, section A4.2). As discussed in chapter 4, section 4.3.4.5, this biphasic data may be indicative of two rate-determining unfolding processes occurring [13]. In the case of the Y107W and L197Y variants, this could be attributed to the disrupted quaternary structures (see sections 5.2.2.3 and 5.2.2.4). The first thermal denaturation temperature was lower than observed for tetrameric

DHDPS species, suggesting that the quaternary structure contributed to thermal stability. The second thermal denaturation temperature was often similar to the thermal denaturation temperatures observed for tetrameric DHDPS (~60 ºC). To rule out the possibility of the presence of abnormal protein folds, CD spectroscopy was employed and is described in section 5.2.2.6.

The disruption of the quaternary structure of DHDPS through the introduction of amino acid substitutions at positions 107 and 197 severely compromised the thermal stability of DHDPS. This reduction in thermal denaturation temperature may significantly affect the propensity of the variants to form amorphous and β-sheet-specific aggregate. This will be investigated in sections 5.2.2.7 and 5.2.2.8.

5.2.2.6 CIRCULAR DICHROISM SPECTROSCOPY

The secondary structures of DHDPS Y107W, its polyhistidine-tagged variants and Y107F were assessed using CD spectroscopy and the results were compared to those obtained for wild-type DHDPS. Figures 5.15 and 5.16 shows that the wavelength spectra of the variants were consistent with those of $(\alpha/\beta)_8$ barrels. Neither the presence of the amino acid substitution at position 107 nor the combinatorial affects of the polyhistidine tags seems to have significantly affected the secondary structure of DHDPS in solution. All the spectra exhibited double minima at 208 nm and 222 nm, characteristic of $(\alpha/\beta)_8$ barrels [14-16] and consistent with the spectra previously obtained for wild-type DHDPS [1, 2, 17]; however, the mean residue ellipticity appears to be slightly lower than wild-type at lower wavelengths. This may indicate that the variants have a more disordered structure than the wild-type enzyme. The polyhistidine tag does not seem to influence this disorder; with the DHDPS Y107W polyhistidine-tagged variants exhibiting similar profiles to non-tagged Y107W, corroborating the results of the crystallographic analysis.

A similar analysis of secondary structure was carried out on pET M11 DHDPS L197Y, pET 151/D-TOPO DHDPS L197Y and cleaved pET 151/D-TOPO DHDPS L197Y (see figure 5.16). The results were compared to those of wild-type DHDPS and demonstrated that although the wavelength scans are similar (each exhibiting double minima at 208 and 222 nm characteristic of an $(\alpha/\beta)_8$ barrel), the amino acid substitution at position 197 also appears to cause a slight change in the secondary structure of DHDPS. The mean residue ellipticity of the DHDPS L197Y variants, like the

Y107 variants, is slightly lower than wild-type at lower wavelengths. This may suggest that the DHDPS L197Y variants also contain more disordered structure than wild-type, contradicting the results reported by Griffin *et al.* [1]. It is possible that in the case of L197Y, the polyhistidine motif is causing the increase in disorder.

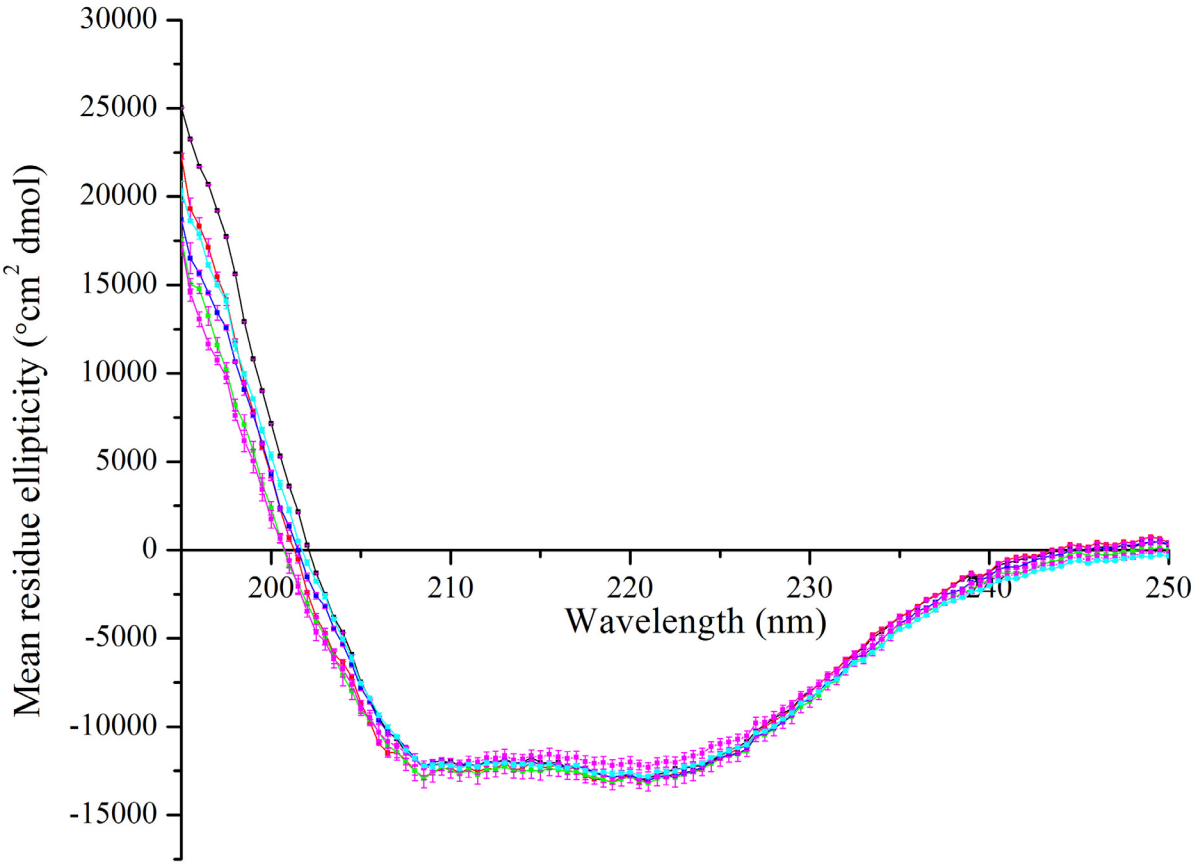


Figure 5.15 – Wavelength scans of wild-type DHDPS (black), DHDPS Y107W (red), pET M11 DHDPS Y107W (green), pET 151/D-TOPO DHDPS Y107W (dark blue), cleaved pET 151/D-TOPO DHDPS Y107W (pink) and DHDPS Y107F (light blue) in 20 mM phosphate buffer (pH 7) containing 40 mM NaCl at 0.2 mg/mL. The plot shows mean residue ellipticity [θ] vs. wavelength ± SEM (N = 3).

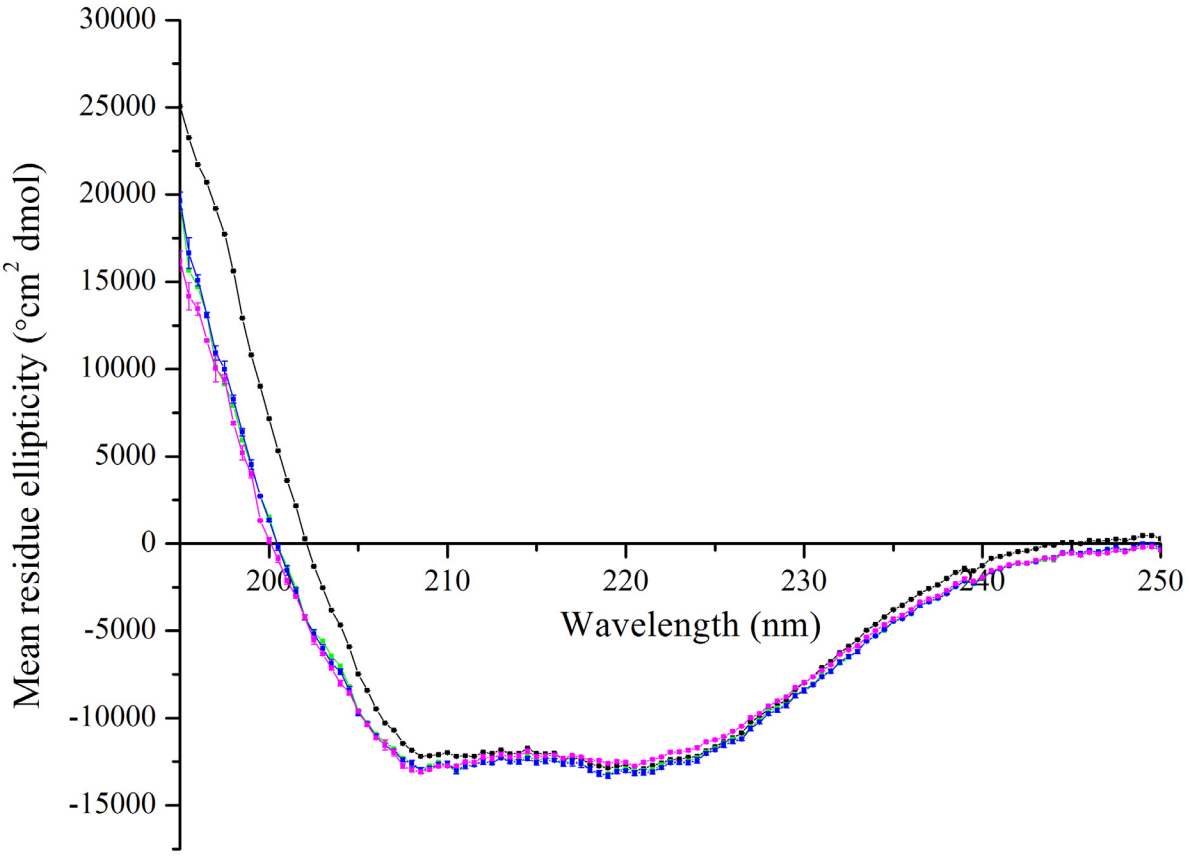


Figure 5.16 – Wavelength scans of wild-type DHDPS (black), pET M11 DHDPS L197Y (green), pET 151/D-TOPO DHDPS L197Y (blue) and cleaved pET 151/D-TOPO DHDPS L197Y (pink) in 20 mM phosphate buffer (pH 7) containing 40 mM NaCl at 0.2 mg/mL. The plot shows mean residue ellipticity vs. wavelength ± SEM (N = 3).

Thermal denaturation experiments were also performed using the CD spectrophotometer by monitoring the change in intensity at 220 nm. All the variants demonstrated thermal denaturation profiles similar to wild-type DHDPS (see appendix 4, figures A4.3 and A4.4). However, the temperatures at which the DHDPS Y107W and L197Y variants thermally denatured were significantly lower than wild-type DHDPS (see table 5.3). The thermal denaturation temperature of DHDPS Y107F was similar to wild-type. The results of this analysis corroborated the results of the DSF and showed good agreement with published literature (in which the thermal denaturation temperature of DHDPS Y107W was 47 ºC by both CD and DSF [2]).

Variant	Thermal denaturation temperature as per CD (°C)	Thermal denaturation temperature as per DSF (°C)
Wild type DHDPS	56.9 ± 0.15	60.90 ± 0.26
DHDPS Y107W	47.4	47.3 ± 0.5
pET M11 DHDPS Y107W	44.0	46.7 ± 0.0
pET 151/D-TOPO DHDPS Y107W	47.5	46.1 ± 0.3
Cleaved pET 151/D-TOPO DHDPS Y107W	47.4	46.1 ± 0.2
DHDPS Y107F	56.2	58.6 ± 0.3
pET M11 DHDPS L197Y	41.8	42.6 ± 0.2
pET 151/D-TOPO DHDPS L197Y	42.7	42.4 ± 0.4
Cleaved pET 151/D-TOPO DHDPS L197Y	48.1	43.2 ± 0.2

Table 5.3 – The thermal denaturation temperatures of DHDPS Y107W, DHDPS L197Y, their polyhistidine-tagged variants and DHDPS Y107F at 0.05 mg/mL in 20 mM phosphate buffer (pH 7, 40 mM NaCl) as monitored by CD spectroscopy and DSF. There are no errors on the CD data as discussed in chapter 2, section 2.3.2.6 and the DSF data are that from the primary thermal denaturation temperature.

The results of the CD spectroscopy demonstrated that DHDPS Y107F, Y107W, L197Y and the polyhistidine-tagged variants of the latter had slightly different wavelength scan profiles compared to wild-type. All the proteins with a disrupted quaternary structure exhibited lower thermal denaturation temperatures than wild-type DHDPS. The effects of the decreased thermal stability on the propensity of the proteins to form both amorphous and β-sheet-specific aggregate are discussed in the following two sections.

5.2.2.7 AMORPHOUS AGGREGATION

The aggregation propensities of DHDPS Y107W, L197Y, their polyhistidine-tagged variants and DHDPS Y107F were screened under all conditions (see chapter 2, section 2.2). The data obtained were compared to the results for wild-type DHDPS. Each assay was carried out as described in chapter 8, section 8.6.5. Although the thermal denaturation temperatures were lower than other DHDPS variants, the assays were still preformed at 60 ºC. This allowed direct comparison to the results obtained for the wild-type protein. However, as the quaternary structure mutants thermally denatured at lower temperatures, the rates of aggregation were expected to be significantly faster than tetrameric DHDPS variants.

The results of the assays are summarised in figures 5.17 – 5.20 and in appendix 4, section A4.4. As described in previous chapters, no increase in light scattering was observed at pH 2 or pH 11. Aggregation at pH 4.5 was the most rapid, followed by pH 7 and pH 9.5 (see figure 5.18 and 5.20). Statistical analysis of both $Agg_{½}$ and the maximum light scattering, was carried out as described in chapter 3, section 3.3.3.6. Significance is presented to a P value of < 0.01. The analysis of deviance for $Agg_{½}$ is presented in appendix 4, table A4.8. The analysis of deviance for maximum absorbance is presented in appendix 4, table A4.9.

The results of the aggregation assays indicated that the presence of the tryptophan at position 107 generally increased the speed of aggregation (see figure 5.17). Where quantifiable, these increases were significant. At pH 7, the addition of the polyhistidine motifs did not appear to affect the half life of aggregation; however, at pH 9.5 they decreased the $Agg_{½}$. The sequence of the polyhistidine motif does not appear to affect the kinetics of aggregation of DHDPS Y107W. The maximum light scattering for all the Y107W variants followed the same pattern as the variants described in previous chapters, exhibiting the highest maximum aggregation at pH 4.5 followed by pH 7 and pH 9.5. Like wild-type, salt does not play a significant role in determining either the $Agg_{½}$ or the maximum light scattering of DHDPS Y107W and its variants (see figure 5.18). At pH 7, the maximum light scattering of DHDPS Y107W was significantly less than wild type. However, the polyhistidine-tagged variants exhibited light scattering similar to that observed for wild-type. The underlying reasons for this variability in maximum light scattering are discussed in chapter 3, section 3.3.3.6. Y107F exhibited slightly elevated rates of thermal denaturation under most conditions; however, usually the maximum light scattering was less than wild-type.

The results of the aggregation assays indicated that the amino acid substitution at position 197 also increased the speed of aggregation compared to wild-type (see figure 5.19). As observed with DHDPS Y107W, the different polyhistidine motifs did not appear to significantly affect the aggregation propensity of the L197Y variants, both exhibiting similar aggregation half lives and maximum aggregation values. The polyhistidine-tagged variants of DHDPS L197Y demonstrated a reduced sensitivity to salt. Where wild-type exhibited some variability in $Agg_{½}$ and maximum light scattering across salt concentrations (see figure 5.20), the effect of salt on the DHDPS L197Y variants was not as significant. The maximum light scattering for the DHDPS L197Y variants followed the same pattern as all other variants, exhibiting the highest maximum aggregation at pH 4.5 followed by pH 7 and pH 9.5. Like other DHDPS variants, salt does not play a significant role in

determining the maximum light scattering except for in the presence of 500 mM NaCl (see figure 5.21).

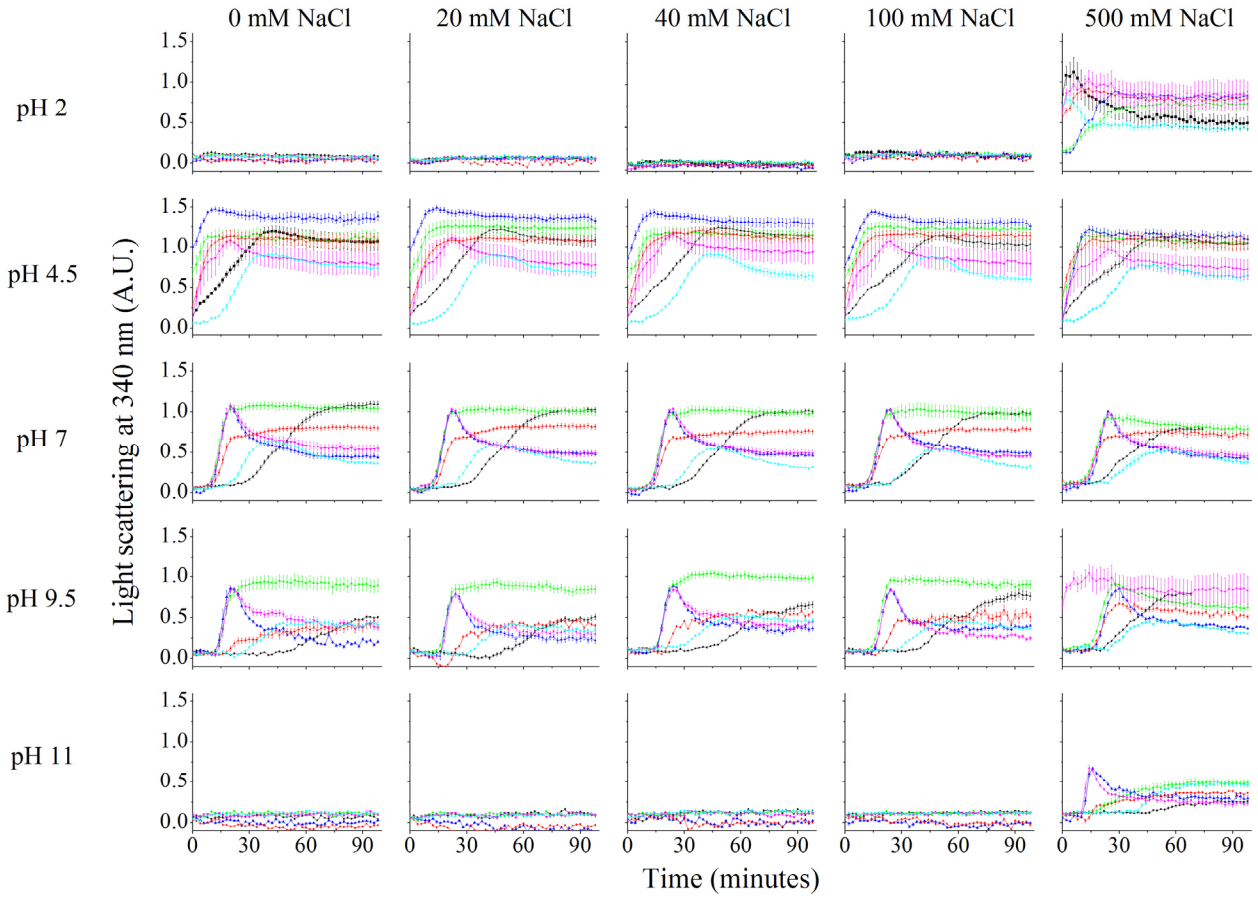


Figure 5.17 – The aggregation propensity of wild-type DHDPS (black), DHDPS Y107W (red), pET M11 DHDPS Y107W (green), pET 151/D-TOPO DHDPS Y107W (dark blue), cleaved pET 151/D-TOPO DHDPS Y107W (pink) and DHDPS Y107F (light blue) in 100 mM phosphate buffer. Where possible, a sigmoid (Boltzmann) function was used for the line of best fit (> 98% confidence interval) in order to calculate the aggregation half life ($Agg_{½}$). The maximum fluorescence was the highest reading over the first 90 minutes of the assay. The data plotted are the mean values of six replicates and the error bars represent the SEM.

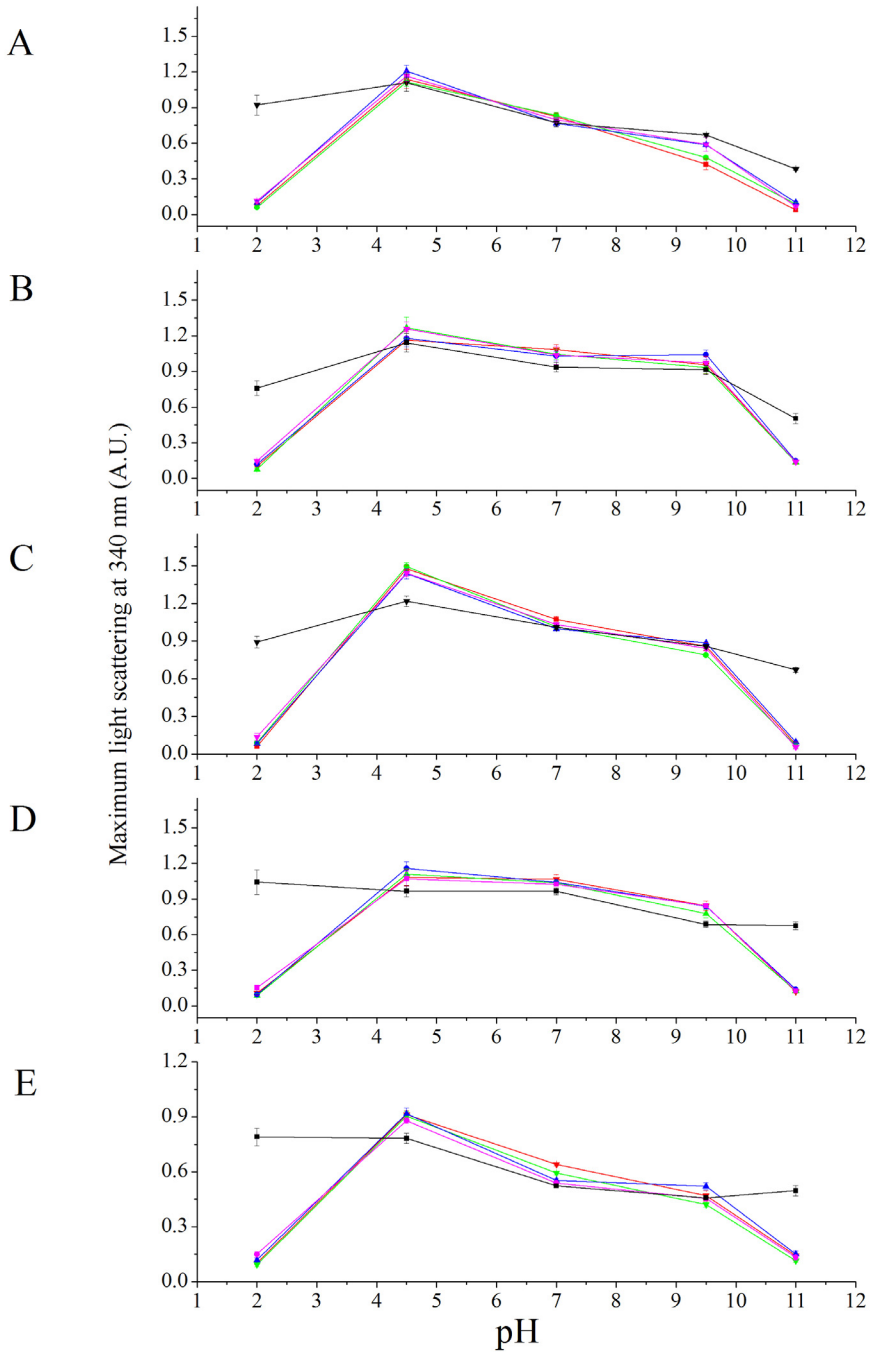


Figure 5.18 – The maximum absorbance of **(A)** DHDPS Y107W, **(B)** pET M11 DHDPS Y107W, **(C)** pET 151/D-TOPO DHDPS Y107W, **(D)** cleaved pET 151/D-TOPO DHDPS Y107W and **(E)** DHDPS Y107F in 100 mM phosphate buffer 0 mM NaCl (red), 20 mM NaCl (green), 40 mM NaCl (blue), 100 mM NaCl (pink) and 500 mM NaCl (black). The data plotted are the maximum absorbance units reached over 90 minutes and are plotted as a function of pH. The data plotted are the mean values of six replicates and the error bars represent the SEM.

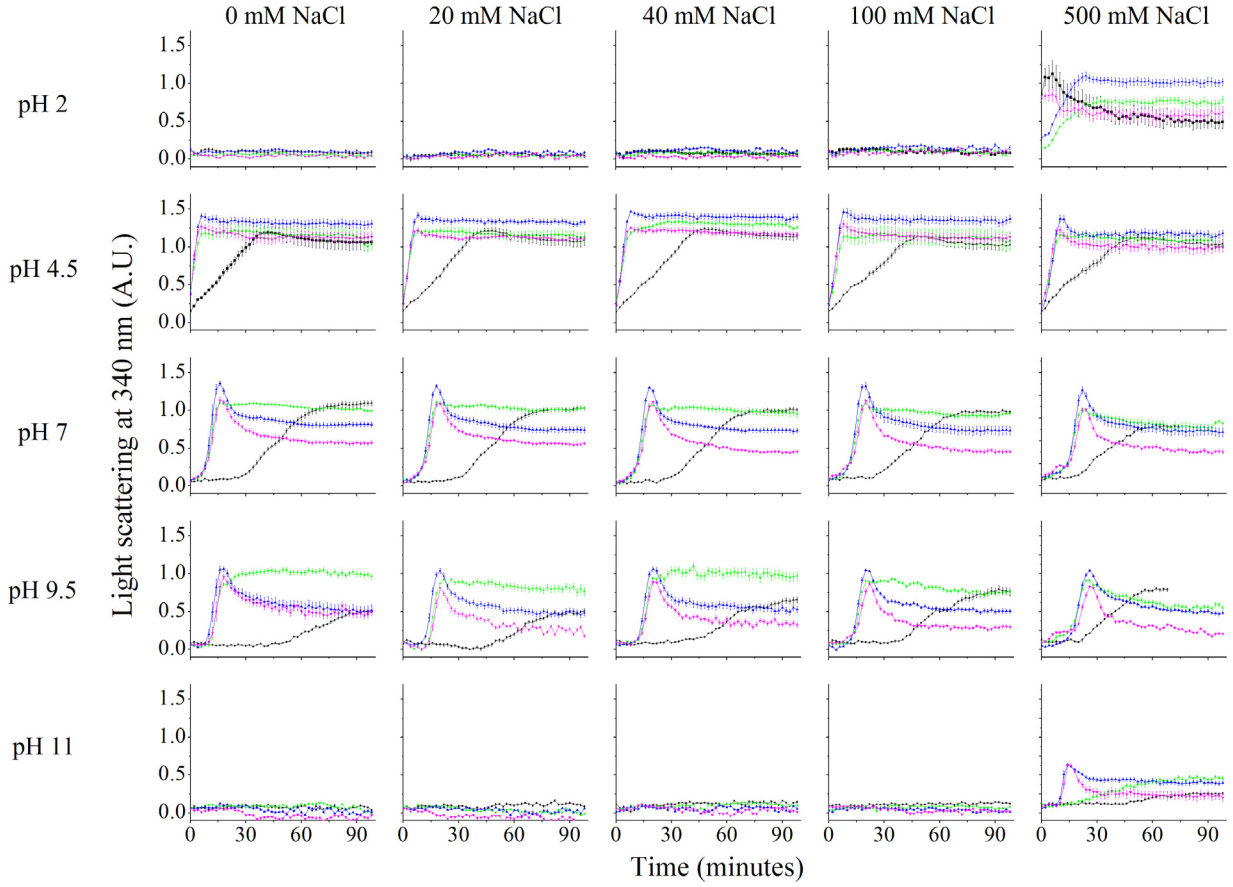


Figure 5.19 – The aggregation propensity of wild-type DHDPS (black), pET M11 DHDPS L197Y (green), pET 151/D-TOPO DHDPS L197Y (blue) and cleaved pET 151/D-TOPO DHDPS L197Y (pink) in 100 mM phosphate buffer. Where possible, a sigmoid (Boltzmann) function was used for the line of best fit (> 98% confidence interval) in order to calculate the aggregation half life ($Agg_{½}$). The maximum fluorescence was the highest reading over the first 90 minutes of the assay. The data plotted are the mean values of six replicates and the error bars represent the SEM.

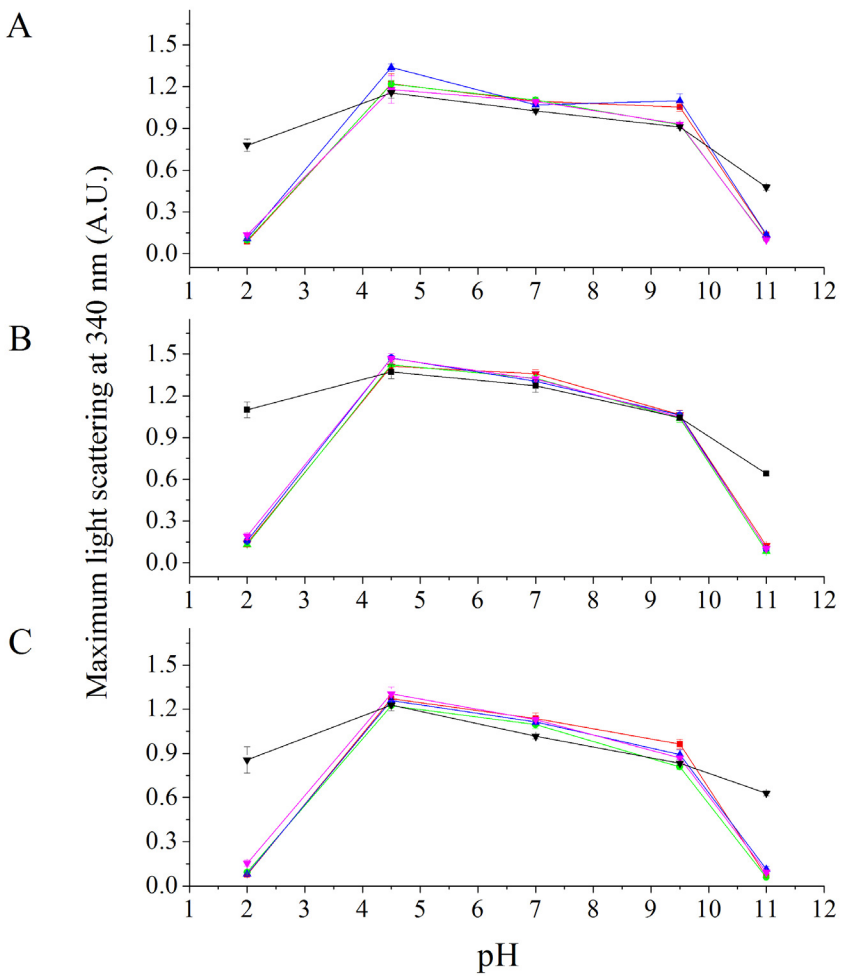


Figure 5.20 – The maximum absorbance of **(A)** pET M11 DHDPS L197Y, **(B)** pET 151/D-TOPO DHDPS L197Y and **(C)** cleaved pET 151/D-TOPO DHDPS L197Y in 100 mM phosphate buffer 0 mM NaCl (red), 20 mM NaCl (green), 40 mM NaCl (blue), 100 mM NaCl (pink) and 500 mM NaCl (black). The data plotted are the maximum absorbance units reached over 90 minutes and are plotted as a function of pH. The data plotted are the means values of six replicates and the error bars represent the SEM.

As described for all other DHDPS variants, the change in amorphous aggregation of DHDPS Y107W, L197Y and their polyhistidine-tagged variants was primarily dependent primarily on pH. As expected, the disruption of the quaternary structure significantly increased the speed of aggregation, either a result of the exposed hydrophobic patches on the surface of the protein or the destabilisation of tertiary structure. The altered propensity to form amorphous aggregate along with the decreased thermal denaturation temperatures (established by DSF and CD) suggested that the amino acid substitutions that disrupt the quaternary structure decreased the stability of DHDPS. It

was postulated that the unstable nature of these variants and altered rate of aggregation would affect the ability of DHDPS to form β-sheet-specific aggregate. This was investigated and is discussed in section 5.2.2.8.

5.2.2.8 β-SHEET-SPECIFIC AGGREGATION

The propensities of DHDPS Y107W, L197Y, their polyhistidine-tagged variants and Y107F to form β-sheet rich aggregate were tested under a range of conditions. Each assay was carried out as described chapter 8, section 8.6.6 and the results were compared to those obtained for wild-type DHDPS. Statistical analysis of $\beta\text{-}agg_{½}$ and the maximum fluorescence (measured as described in chapter 2, section 2.3.2.8) were carried out as described in chapter 3, section 3.3.3.7. Significance is presented to a P value of < 0.01. The analysis of deviance for $\beta\text{-}agg_{½}$ is presented in appendix 4, table A4.11. The analysis of deviance for maximum absorbance is presented in appendix 4, table A4.12.

As discussed in chapter 2, section 2.3.2.8, the only comparison possible was between the protein variants and wild-type due to the effects of salt and pH on the kinetics of β-sheet-specific aggregation and ThT binding, and the effect of pH on the affinity of ThT for β-sheet structures. This comparative analysis of the protein variants revealed that the polyhistidine tags played the primary role in determining the maximum fluorescence (see figure 5.21). The fluorescence of DHDPS Y107W and Y107F remained similar to the wild-type whereas the polyhistidine-tagged variants demonstrated a significant reduction in $\beta\text{-}agg_{½}$ and a significant increase in maximum fluorescence. This result corroborates the results described in chapter 3, section 3.3.3.7 and chapter 4, section 4.3.4.8. The pET 151/D-TOPO polyhistidine tag resulted in the greatest increase in fluorescence with the cleaved variant also exhibiting higher fluorescence than the wild-type. These results differ from those described in chapter 4 and 5 which found the pET M11 polyhistidine tag to cause the greatest increase in fluorescence. The Y107F amino acid substitution did not significantly increase the propensity of DHDPS to form β-sheet-specific aggregate (see figure 5.21).

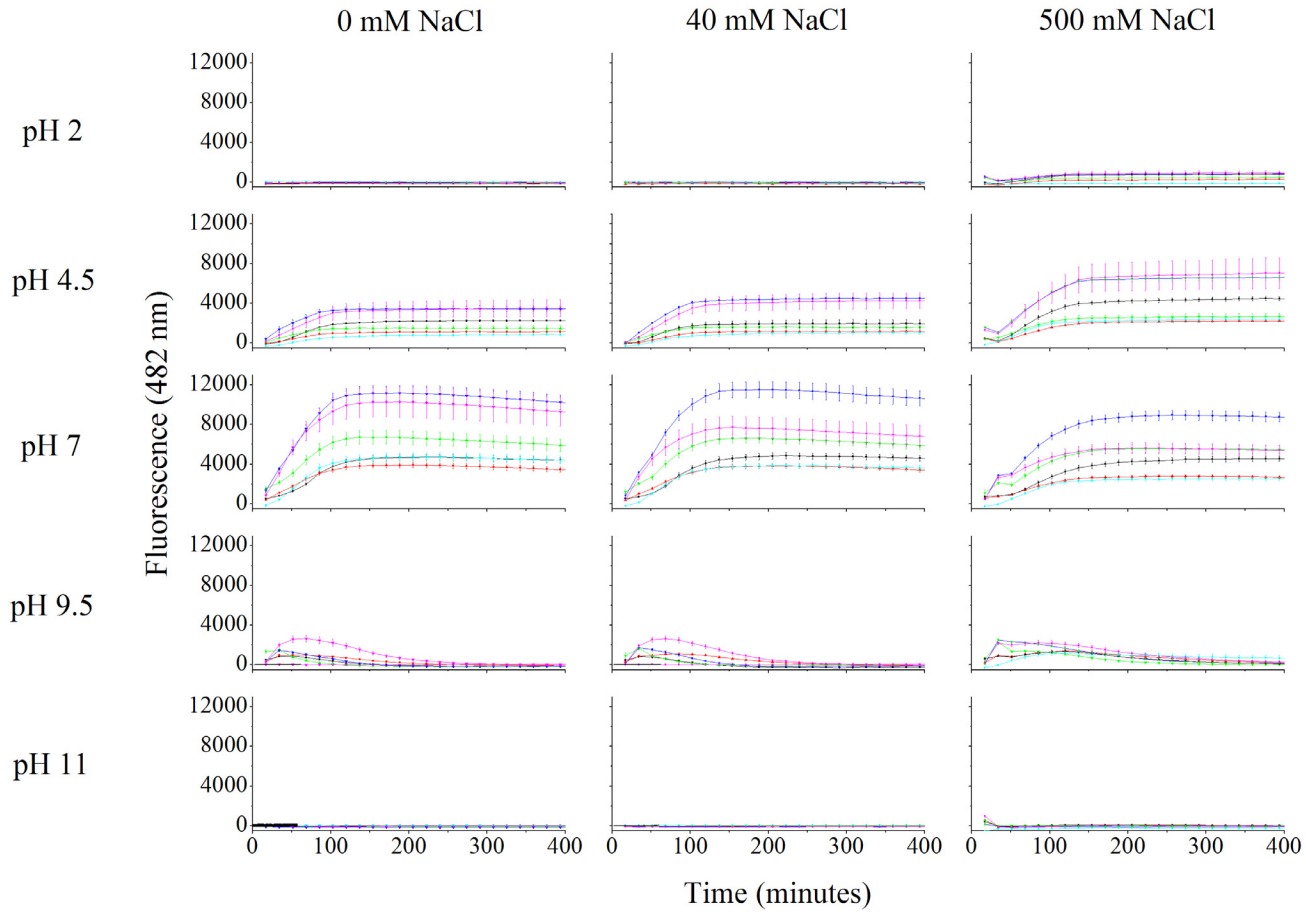


Figure 5.21 – The ThT fluorescence in the presence of wild-type DHDPS (black), DHDPS Y107W (red), pET M11 DHDPS Y107W (green), pET 151/D-TOPO DHDPS Y107W (dark blue), cleaved pET 151/D-TOPO DHDPS Y107W (pink) and DHDPS Y107F (light blue) in 100 mM phosphate buffer. The data plotted are the mean values of six replicates ± SEM. The β-$agg_{½}$ was determined by identifying the midpoint of a fitted sigmoid curve as previously described [18]. The maximum fluorescence was the highest reading over the first 300 minutes of the assay.

The pattern observed for the polyhistidine-tagged variants of Y107W was similar to the DHDPS L197Y variants which also demonstrated decreased β-$agg_{½}$ and increased maximum fluorescence for the pET 151/D-TOPO polyhistidine tag and its cleaved counterpart, compared to wild-type (see figure 5.22). However, the presence of the amino acid substitution at position 197 with the pET M11 polyhistidine tag reduced the maximum fluorescence. This result contradicts results described in the previous chapters which found the pET M11 polyhistidine motif to increase the propensity of DHDPS to form β-sheet-specific aggregate significantly.

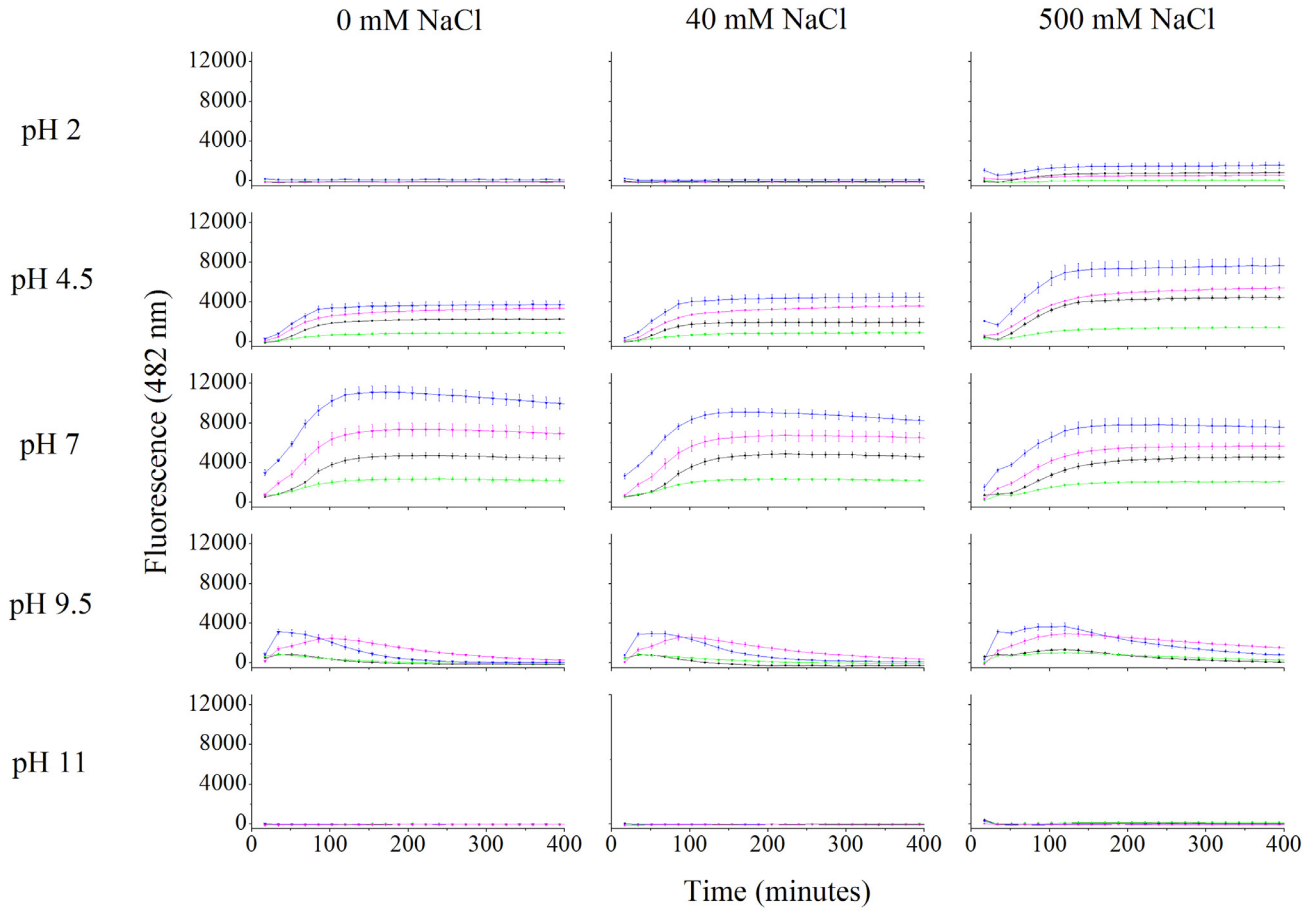


Figure 5.22 – The ThT fluorescence in the presence of wild-type DHDPS (black), pET M11 DHDPS L197Y (green), pET 151/D-TOPO DHDPS L197Y (blue) and cleaved pET 151/D-TOPO DHDPS L197Y (pink) in 100 mM phosphate buffer. The data plotted are the mean values of six replicates ± SEM. The β-$agg_{½}$ was determined by identifying the midpoint of a fitted sigmoid as previously described [18]. The maximum fluorescence was the highest reading over the first 300 minutes of the assay.

Despite the inability to directly compare the effects of buffer condition on the fluorescence of the DHDPS variants in the presence of ThT, general trends can be observed. This indicates that the fluorescence of ThT is highest at pH 7 for both the Y107W and L197Y variants (see figure 5.23 and 5.23).

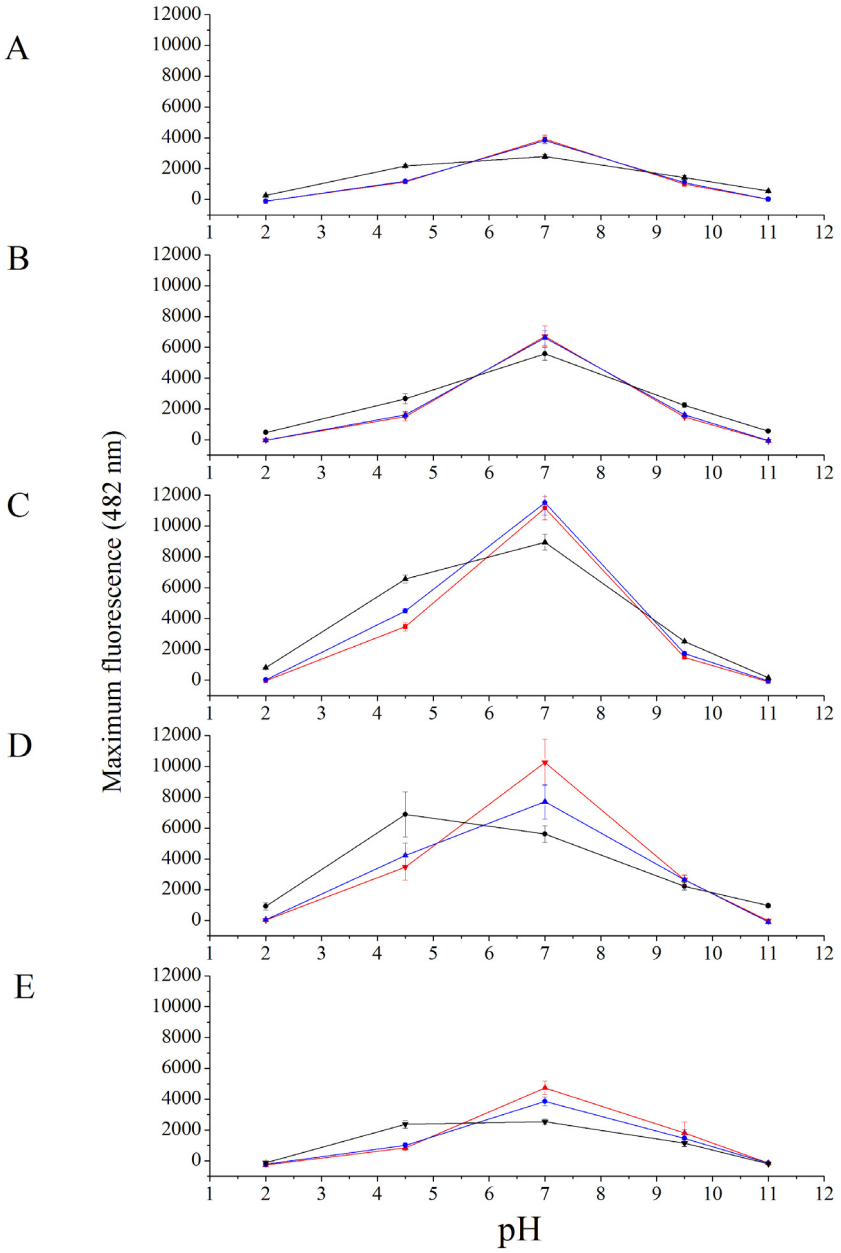


Figure 5.23 – The maximum fluorescence of **(A)** DHDPS Y107W, **(B)** pET M11 DHDPS Y107W, **(C)** pET 151/D-TOPO DHDPS Y107W, **(D)** cleaved pET 151/D-TOPO DHDPS Y107W and **(E)** DHDPS Y107F in 100 mM phosphate buffer, 0 mM NaCl (red), 40 mM NaCl (blue) and 500 mM NaCl (black). The data plotted are the maximum fluorescence reached over the initial 300 minutes of the assay and are plotted as a function of pH. The data plotted are the mean values of six replicates and the error bars represent the SEM.

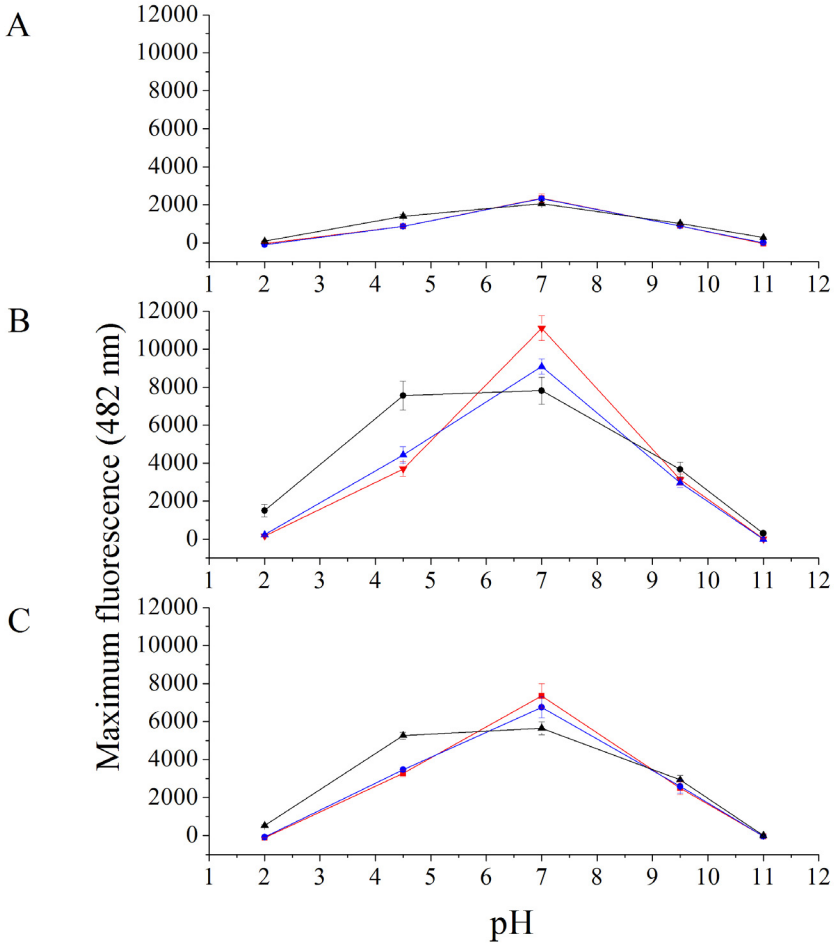


Figure 5.24 – The maximum fluorescence of **(A)** pET M11 DHDPS L197Y, **(B)** pET 151/D-TOPO DHDPS L197Y and **(C)** cleaved pET 151/D-TOPO DHDPS L197Y in 100 mM phosphate buffer, 0 mM NaCl (red), 40 mM NaCl (blue) and 500 mM NaCl (black). The data plotted are the maximum fluorescence reached over the initial 300 minutes of the assay and are plotted as a function of pH. The data plotted are the mean values of six replicates and the error bars represent the SEM.

The combinatorial effects of the amino acid substitutions at position 107 and 197 and the extra amino acids associated with the polyhistidine motifs on the ThT fluorescence appeared to indicate that the polyhistidine motifs exerted the greatest influence on the formation of β-sheet-specific aggregate. The effects of the amino acid substitutions that disrupt quaternary structure on the stability of DHDPS did not seem to influence the β-sheet-specific aggregation. However, the decreased stability may have been indicative of changes in misfolding mechanisms. Thus the samples exhibiting the highest ThT fluorescence were further tested for the presence of amyloid. The results of these assays are described in the following section (section 5.2.2.9)

5.2.2.9 CONFIRMATION OF AMYLOID FIBRILS

Confirmation of the presence of amyloid fibrils by TEM was attempted for the DHDPS Y107W and L197Y polyhistidine-tagged variants. The samples were prepared under the conditions which exhibited the highest fluorescence determined in section 5.2.2.8 and as described in chapter 8, section 8.8.1. Fibrillar structures were observed in the pET 151/D-TOPO DHDPS Y107W samples by TEM. These are shown in figure 5.25. The fibril-like structures were mixed with amorphous aggregate. If these linear aggregates are amyloid fibrils, these results represent the first reported occurrence of an $(\alpha/\beta)_8$ barrel forming amyloid fibrils.

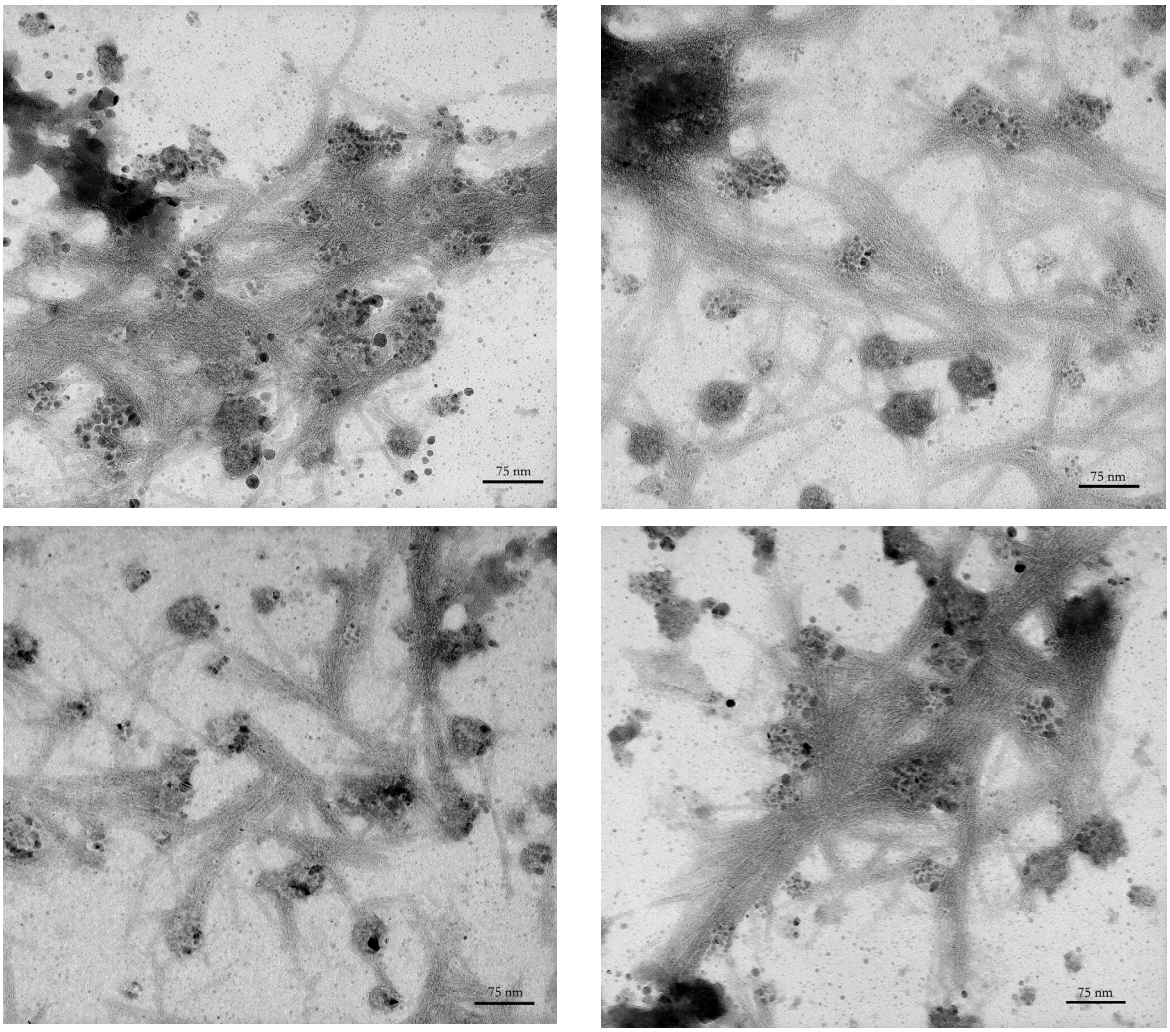


Figure 5.25 – TEM image of linear aggregate formed by pET 151/D-TOPO DHDPS Y107W in 100 mM phosphate buffer (pH 7), 40 mM NaCl at 60 ºC. This image was typical of all 15 images obtained.

As not all fibrillar aggregates can be considered amyloid fibrils [19], attempts were made to confirm their presence through X-ray fibre diffraction. Stalks were prepared as described in chapter 8, section 8.8.2. The presence of amorphous aggregate in the sample made the preparation difficult. Attempts were made to isolate the fibrillar material though centrifugation and protease treatment. However, following these treatments the linear aggregates could no longer be identified by TEM and no longer fluoresced in the presence of ThT.

5.3 SUMMARY AND CONCLUSIONS

As per previous studies [1, 2, 6] it was shown that the disruption of the quaternary structure of DHDPS has a significant effect on the catalytic activity of the enzyme. Interestingly, the presence of the polyhistidine tags on Y107W appeared to result in the formation of an entirely monomeric species as judged by gel filtration; however, this did not affect the catalysis further, most likely due to the fact that the polyhistidine-tagged Y107W variants existed in monomer/tetramer equilibrium in solution as judged by AUC. The substitution of tyrosine at 107 for a phenylalanine resulted in attenuated activity due to its involvement in catalysis by DHDPS [5]. It has been hypothesised that the quaternary structure confers improved catalytic activity by minimising the movement between the tight dimers (see figure 5.1) [1]. From the results of the DSF and the CD analysis of the thermal stability of the variants, it appears that the tetrameric structure also maintains the thermal stability of DHDPS. Those with aberrant quaternary structures exhibited significantly decreased thermal denaturation temperatures compared to wild-type and DHDPS Y107F. No significant changes in the secondary or tertiary structure were seen in the crystal structures which supports the hypothesis that the destabilisation of the proteins as assessed by DSF and CD was due to the disrupted quaternary structures as assessed by gel filtration and AUC. This destabilisation is again apparent in the propensity of the *E. coli* DHDPS Y107W, L197Y and their polyhistidine-tagged variants to form amorphous aggregate. The factor with the most influence on the rate and extent of aggregate formed is the disrupted quaternary structures. The presence of the polyhistidine tags plays a secondary role in influencing aggregation in the case of these point mutations. As per previously described for proteins containing other amino acid substitutions (see chapters 3, 4 and 5), pH has a considerable impact on aggregation.

The incubation of the proteins with ThT resulted in some very significant fluorescence readings. Contrary to the results obtained from monitoring the amorphous aggregation, the presence of the polyhistidine tags or the amino acids remaining following cleavage of the pET 151/D-TOPO polyhistidine tag had a significant influence on the fluorescence. The fluorescence in the presence of thioflavin T indicated that β-sheet-specific aggregate formation was determined primarily by the presence of the polyhistidine tags and suggested that β-sheet-specific aggregate was formed. TEM analysis demonstrated the presence of fibrillar structures. However, confirmation of this by X-ray fibre diffraction was not achieved.

The data presented in this chapter suggest that the combinatorial affect of amino acids that disrupt the quaternary structure of DHDPS and polyhistidine motifs may have resulted in the formation of a DHDPS species prone to forming linear aggregates. It is possible that this is due to the increased rate of aggregation facilitating the formation of an intermediate species which enabled the formation of amyloid fibrils. Further confirmation is required before the linear aggregates can be definitively identified as amyloid fibrils; however, if they are amyloid, this would be the first report of an $(\alpha/\beta)_8$ barrel forming fibrils. Moreover, the results in this chapter strongly support the hypothesis that the quaternary structure of DHDPS affords protection against aggregation.

5.4 REFERENCES

1 Griffin, M. D., Dobson, R. C. J., Pearce, F. G., Antonio, L., Whitten, A. E., Liew, C. K., Mackay, J. P., Trewhella, J., Jameson, G. B., Perugini, M. A. and Gerrard, J. A. (2008) Evolution of quaternary structure in a homotetrameric enzyme. Journal of Molecular Biology **380**, 691 - 703

2 Pearce, F. G., Dobson, R. C., Weber, A., Lane, L. A., McCammon, M. G., Squire, M. A., Perugini, M. A., Jameson, G. B., Robinson, C. V. and Gerrard, J. A. (2008) Mutating the tight-dimer interface of dihydrodipicolinate synthase disrupts the enzyme quaternary structure: towards a monomeric enzyme. Biochemistry **In press**

3 DeLano, W. L. (2002) The PyMOL molecular graphics system. DeLano Scientific, San Carlos

4 Dobson, R. C. J., Griffin, M. D. W., Jameson, G. B. and Gerrard, J. A. (2005) The crystal structures of native and (*S*)-lysine-bound dihydrodipicolinate synthase from *Escherichia coli* with improved resolution show new features of biological significance. Acta Crystallographica. Section D, Biological Crystallography **D61**, 1116 - 1124

5 Dobson, R. C. J., Valegard, K. and Gerrard, J. A. (2004) The crystal structure of three site-directed mutants of *Escherichia coli* dihydrodipicolinate synthase: Further evidence for a catalytic triad. Journal of Molecular Biology **338**, 329 - 339

6 Griffin, M. D. W. (2005) PhD Thesis. Why is DHDPS a tetramer?, University of Canterbury, Christchurch

7 Andrews, P. (1964) Estimation of the molecular weights of proteins by Sephadex gel filtration. Biochemical Journal **91**, 222 - 233

8 Andrews, P. (1965) The gel filtration behaviour of proteins related to their molecular weights over a wide range. Biochemical Journal **96**, 595 - 606

9 Perugini, M. A., Griffin, M. D. W., Smith, B. J., Webb, L. E., Davin, A. J., Handman, E. and Gerrard, J. A. (2005) Insight into the self-association of key enzymes from pathogenic species. European Biophysical Journal **34**, 469 - 476

10 Schuck, P., Perugini, M. A., Gonzales, N. R., Howlett, G. J. and Schubert, D. (2002) Size-distribution analysis of proteins by analytical ultracentrifugation and Lamm equation modeling. Biophysical Journal **78**, 1606 - 1619

11 Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. Biophysical Journal **78**, 1606 - 1619

12 Vedadi, M., Niesen, F. H., Allali-Hassani, A., Fedorov, O. Y., Finerty, P. J. J., Wasney, G. A., Yeung, R., Arrowsmith, C., Ball, L. J., Berglund, H., Hui, R., Marsden, B. D., Norlund, P., Sundstrom, M., Weigelt, J. and Edwards, A. M. (2006) Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. Proceedings of the National Academy of Sciences USA **103**, 15835 - 15840

13 Epps, D. E., Sarver, R. W., Rogers, J. M., Herberg, J. T. and Tomich, P. K. (2001) The ligand affinity of proteins measured by isothermal denaturation kinetics. Analytical Biochemistry **292**, 40 - 50

14 Singh, K. and Bhakuni, V. (2007) Cation induced differential effect on structural and functional properties of *Mycobacterium tuberculosis* α-isopropylmalate synthase. BMC Structural Biology **7**, 39 - 49

15 Manavalan, P. and Johnson Jr., W. C. (1983) Sensitivity of circular dichroism to protein tertiary structure class. Nature **305**, 831 - 832

16 Offerdi, F., Dubail, F., Kischel, P., Sarinski, K., Stern, A. S., Weerdt, C. v. d., Hoch, J. C., Prosperi, C., Francois, J. M., Mayo, S. L. and Martial, J. A. (2003) *De novo* backbone and sequence design of an idealized α/β-barrel protein: Evidence of stable tertiary structure. Journal of Molecular Biology **325**, 163 - 174

17 Dobson, R. C. J., Devenish, S. R. A., Turner, L. A., Clifford, V. R., Pearce, F. G., Jameson, G. B. and Gerrard, J. A. (2005) Role of arginine 138 in the catalysis and regulation of *Escherichia coli* dihydrodipicolinate synthase. Biochemistry **44**, 13007 - 13013

18 Sabate, R., Gallardo, M. and Estelrich, J. (2003) An autocatalytic reaction as a model for the kinetics of the aggregation of β-amyloid. Biopolymers (Peptide Science) **71**, 190 - 195

19 Chiti, F. and Dobson, C. M. (2006) Protein misfolding, functional amyloid and human disease. Annual Review of Biochemistry **75**, 333 - 366

CHAPTER 6

COMBINING AN ALGORITHM MUTATION WITH A QUATERNARY STRUCTURE VARIANT

6.1 INTRODUCTION

The predictions of the algorithm Zyggregator are built on the assumption that the protein is monomeric and unfolded. As discussed in chapter 1 and 2, DHDPS is tetrameric and under most experimental conditions exists in a folded state. Therefore, the assumptions of the algorithm were unlikely to be correct. In order to maximize the chance of increasing the yield of amyloid fibrils, a combinatorial approach was taken: the amino acid substitution at position 107 (which disrupts quaternary structure) was combined with Q90L (an amino acid substitution arising from the results of the algorithm). The hypothesis that these two amino acid substitutions would combinatorially increase the aggregation propensity of DHDPS was supported by the results of the Zyggregator algorithm, which indicated that the DHDPS species containing both amino acid substitutions are more prone to forming amyloid fibrils (with Z_{agg} scores of 0.37 for pET 151/D-TOPO DHDPS Q90L/Y107W and 0.57 for cleaved pET 151/D-TOPO DHDPS Q90L/Y107W compared to -0.46 for wild-type, 0.36 for pET 151/D-TOPO DHDPS Q90L and -0.52 for pET 151/D-TOPO DHDPS Y107W). Complete results of the Zyggregator analyses are presented in appendix 5, table A5.1.

Preliminary characterisation of DHDPS Q90L/ Y107W was carried out on the pET 151/D-TOPO variant and its cleaved derivative. As the discussion below indicates, the effects of the two amino acid substitutions on aggregation propensity were not additive. For this reason the characterisation of the pET M11 variant was not pursued.

6.2 RESULTS

6.2.1 PURIFICATION

The pET 151/D-TOPO DHDPS Q90L/Y107W was purified by affinity chromatography, as described in chapter 8, sections 8.4.1, 8.4.4, and 8.8.7 - 8.4.9. The yield averaged 30.7 mg (0.68 units with specific activity of 0.02 units/mg) per litre of bacterial culture. This compared to the yield of wild-type which averaged 19.6 mg (76 units with a specific activity of 3.83 units/mg) per litre of bacterial culture, pET 151/D-TOPO DHDPS Q90L which averaged yielded an average 39 mg (126 units with a specific activity of 3.26 units/mg) per litre of culture, and pET 151/D-TOPO DHDPS Y107W yielded 40.3 mg (2.4 units with a specific activity of 0.06 units/mg) per litre of bacterial culture. The addition of the polyhistidine tag to DHDPS Q90L/Y107W improved the total protein yield as described for other DHDPS variants (describe in previous chapters); however, the specific activity was greatly reduced. This was attributed to the presence of the Y107W amino acid substitution which, both disrupts the quaternary structure, and attenuates catalytic function (described in chapter 5, section 5.2.2.1). This hypothesis was supported by SDS-PAGE gel analysis, which demonstrated a high level of purity in the protein sample (see figure 6.1).

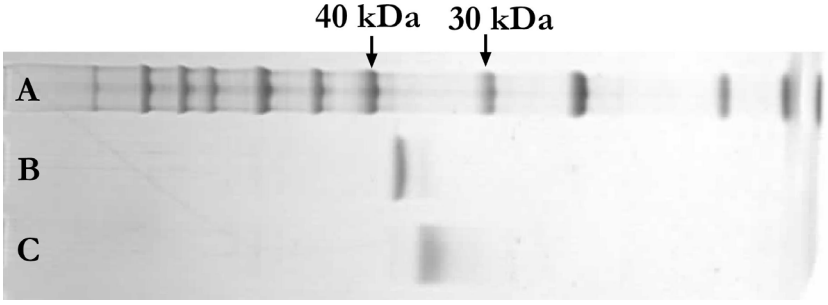


Figure 6.1 – SDS-PAGE. The lanes were loaded with **(A)** molecular weight marker with relevant bands indicated, **(B)** pET 151-D-TOPO DHDPS Q90L/Y107W and **(C)** cleaved pET 151-D-TOPO DHDPS Q90L/Y107W.

The yield following cleavage of the polyhistidine motif (using the technique described in chapter 8, section 8.4.9) from tagged DHDPS Q90L/Y107W, averaged 56 % of the original protein (17 mg (0.83 units with specific activity of 0.05 units/mg) per litre of bacterial culture). These results suggest

that upon removal of the polyhistidine motif, the catalytic efficiency is marginally improved; however, the activity was still severely attenuated compared to the wild-type enzyme. Again, this was attributed to the importance of Y107 in the catalytic triad [1].

6.2.2 BIOPHYSICAL PROPERTIES

The suite of techniques described in chapter 2 was used to characterise pET 151/D-TOPO DHDPS Q90L/Y107W and its cleaved variant. The results of these assays are described in this chapter and for each assay, the results were compared to those obtained for wild-type DHDPS, pET 151/D-TOPO DHDPS Q90L and pET 151/D-TOPO DHDPS Y107W.

6.2.2.1 KINETICS

pET M11 DHDPS Q90L/Y107W and its cleaved variant were assessed for their catalytic ability as described in earlier chapters. The results of this analysis revealed that the catalytic capacity of the DHDPS Q90L/Y107W variants was seriously attenuated (table 6.1). This is a similar result to those presented in chapter 5, section 5.2.2.1 which demonstrated that the disruption of the quaternary structure resulted in a significant decrease in catalytic efficacy.

	k_{cat}^{app} with respect to Pyruvate (s^{-1})	$K_{m\ Pyr}^{app}$ (mM)	k_{cat}^{app} with respect to (*S*)-ASA (s^{-1})	$K_{m\ (S)\text{-}ASA}^{app}$ (mM)
Wild-type DHDPS	118 ± 3	0.3 ± 0.03	123 ± 5	0.14 ± 0.02
pET 151/D-TOPO DHDPS Q90L	101 ± 5	0.54 ± 0.10	102 ± 6	0.44 ± 0.07
pET 151/D-TOPO DHDPS Y107W	1.11 ± 0.10	6.43 ± 0.15	2.65 ± 0.18	0.61 ± 0.11
pET 151/D-TOPO DHDPS Q90L/Y107W	0.67 ± 0.02	1.61 ± 0.13	0.72 ± 0.03	0.40 ± 0.05
Cleaved pET 151/D-TOPO DHDPS Q90L/ Y107W	0.59 ± 0.03	1.59 ± 0.29	2.42 ± 0.11	0.42 ± 0.05

Table 6.1 –The kinetic parameters for the pET M11 DHDPS Q90L/Y107W and cleaved pET 151/D-TOPO DHDPS Q90L/Y107W compared to those of the parental mutants and wild-type. The two different k_{cat} values arise from the use of pseudo-single substrate models rather than a complete two substrate analysis.

The apparent Michaelis constants for (*S*)-ASA and pyruvate were increased for both double mutant variants compared to wild-type. These results are similar to those of the Y107W variants (described in chapter 5, section 5.2.2.1) and confirmed that disruption of the quaternary structure due to the presence of the tryptophan at position 107 caused a significant increase in the K_m^{app} for both substrates however the additional substitution at position 90 has reduced the K_m^{app} for both substrates compared to pET 151/D-TOPO DHDPS Y107W.

The k_{cat}^{app} was drastically reduced for both DHDPS Q90L/Y107W variants which is consistent with the results obtained for the mutants with aberrant quaternary structures. From the results it appears that the additive effects of the amino acid substitution at position 90 were minimal, with the major determinant of catalytic ability being the presence of the tryptophan at position 107 in place of the tyrosine.

6.2.2.2 CRYSTALLOGRAPHIC ANALYSIS

As with all previous DHDPS enzymes, attempts to crystallise pET 151/D-TOPO DHDPS Q90L/Y107W and its cleaved variant were unsuccessful.

6.2.2.3 ANALYTICAL GEL PERMEATION LIQUID CHROMOTOGRAPHY

The results of analytical gel permeation liquid chromatography are shown in figure 6.2. The molecular masses of the two variants were compared to the calculated molecular mass, as in the previous chapters and indicated that both double mutants exist primarily as monomers in solution. The disparity between the calculated and predicted molecular masses is outside the accepted 10 % error [2, 3] and may be indicative of the protein acting in a non-ideal manner while on the column.

The cleaved variant of pET 151/D-TOPO DHDPS Q90L/Y107W eluted at a volume consistent with a molecular mass of 35 kDa (see figure 6.2). This indicated that the cleaved double mutant also existed primarily as a monomer in solution, a result consistent with that observed for pET 151/D-TOPO DHDPS Y107W. The monomeric nature of DHDPS Q90L/Y107W is likely to have been bought about through the combinatorial affect of the tryptophan at position 107 and the pET 151/D-TOPO motif. The secondary peak apparent on the chromatogram may be due to the

presence of an alternate oligomeric state, possibly a dimer given its location. This is consistent with the results previously described for DHDPS Y107W variants (see chapter 5, section 5.2.2.3). Although the proteins were all 0.5 mg/mL, variability in the mAU was observed. This may be due to the protein aggregating on the column. This propensity of the double mutant to aggregate is investigated further in section 6.2.2.6.

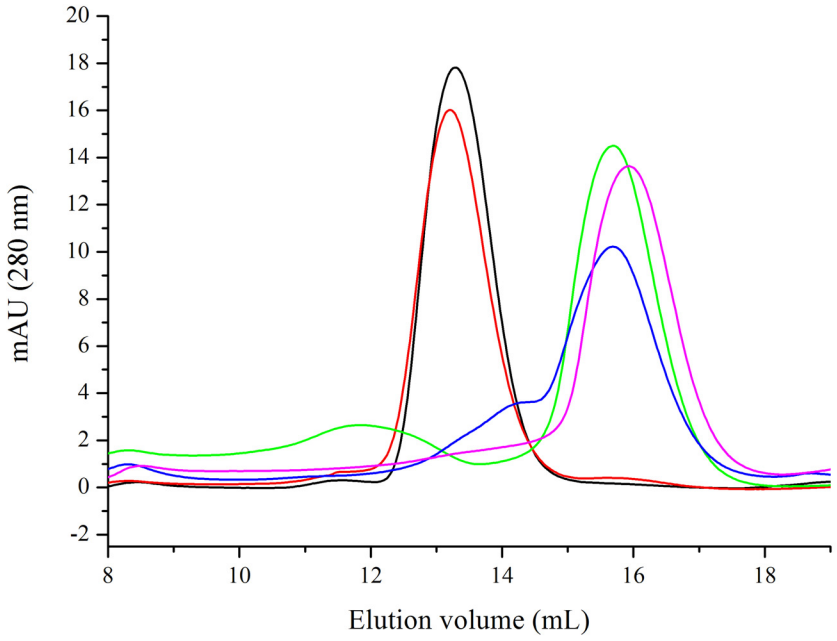


Figure 6.2 – Elution profiles determined by zonal analysis for **(A)** wild-type DHDPS (black), pET 151/D-TOPO DHDPS Q90L (red), pET 151/D-TOPO DHDPS Y107W (green), pET 151/D-TOPO DHDPS Q90L/Y107W (blue), and cleaved pET 151/D-TOPO DHDPS Q90L/Y107W (pink). Data indicate that both the double mutants have aberrant quaternary structures. The chromatograms were produced by measuring absorbance at 280 nm as a function of elution volume (V_e).

6.2.2.4 DIFFERENTIAL SCANNING FLUORIMETRY

The combinatorial effect of the amino acid substitutions at positions 90 and 107, and the presence of the extra-molecular residues associated with the intact and cleaved pET 151/D-TOPO polyhistidine tag on the thermal denaturation temperature of DHDPS were assessed by DSF (shown in figure 6.3 and presented in their entirety in appendix 5, section A5.2). These results indicate that the thermal stability of DHDPS is severely compromised for both the polyhistidine tagged and cleaved variants compared to wild-type DHDPS. Previous results (chapter 3, section 3.3.3.4) suggested that the

presence of the Q90L amino acid substitution only had a small effect on the thermal denaturation temperature of DHDPS. However, the substitution of the tyrosine at position 107 with a tryptophan, results in a significant decrease in thermal denaturation temperature (see chapter 5, section 5.2.2.5). This was attributed to the disruption of the quaternary structure destabilising DHDPS [4, 5]. From these results it appears that the main determinant of the thermal stability is the Y107W amino acid substitution. As described in earlier chapters, the presence of the pET 151/D-TOPO polyhistidine tag or the residues remaining following cleavage do not significantly effect the thermal denaturation temperature of the proteins.

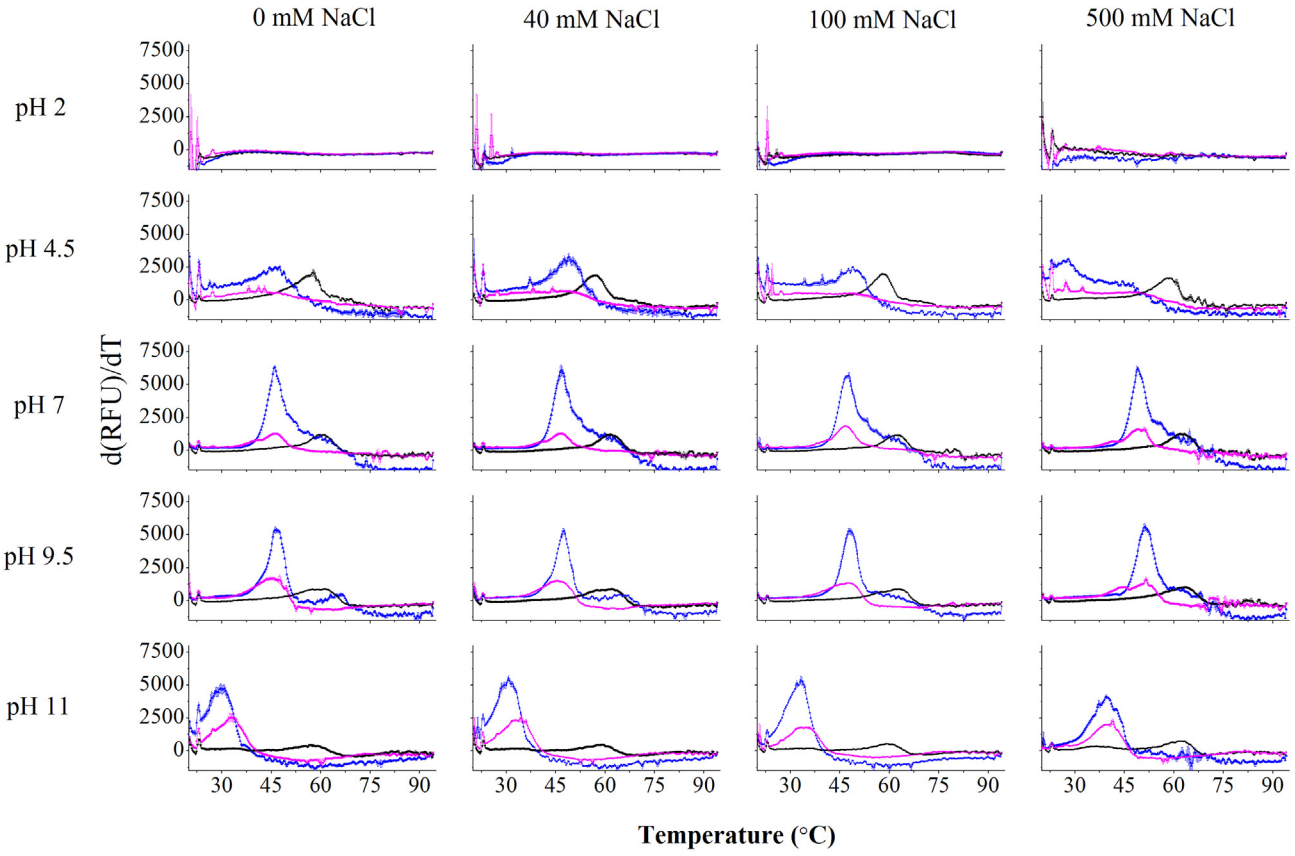


Figure 6.3 – The thermal denaturation profiles of wild-type DHDPS (black), pET 151/D-TOPO DHDPS Q90L/Y107W (blue) and cleaved pET 151/D-TOPO DHDPS Q90L/Y107W (pink) at 0.5 mg/mL in 100 mM phosphate buffer (pH 7, 40 mM NaCl) as monitored by SYPRO orange fluorescence. The data plotted are the derivatives of the increase in fluorescence monitored by the BioRad IQ5 (presented in appendix A5, figure A5.1). The peaks indicate the temperature at which the protein is unfolding the most rapidly and is recorded as the mean thermal denaturation temperature. The data plotted are the mean values of three replicates ± SEM.

The effect of salt on the thermal denaturation temperatures was (as described for other variants) minimal with the only significant differences occurring at pH 4.5, 500 mM NaCl. The effect of pH is more significant. As described in previous chapters, the proteins exhibited high initial fluorescence suggesting that hydrophobic patches were exposed prior to the addition of the dye [6]. Generally, the thermal denaturation temperatures of DHDPS were highest at pH 7 followed by pH 9.5, and exhibited lower melting temperatures at pH 4.5 and pH 11 (see chapter 2, section 2.3.2.5 and chapter 3, section 3.3.3.4). The variants with aberrant quaternary structures demonstrate a decreased sensitivity to pH (see chapter 5, section 5.2.2.5) with the thermal denaturation temperature at pH 4.5 being similar to those at pH 7 and pH 9.5. pET 151/D-TOPO DHDPS Q90L/Y107W exhibited such a trend (see figure 6.4), indicating that the amino acid substitution at position 107 was the major influence on thermal stability. The cleaved variant exhibited a trend analogous to the variants with intact quaternary structure, with higher thermal denaturation temperatures at pH 7 and pH 9.5.

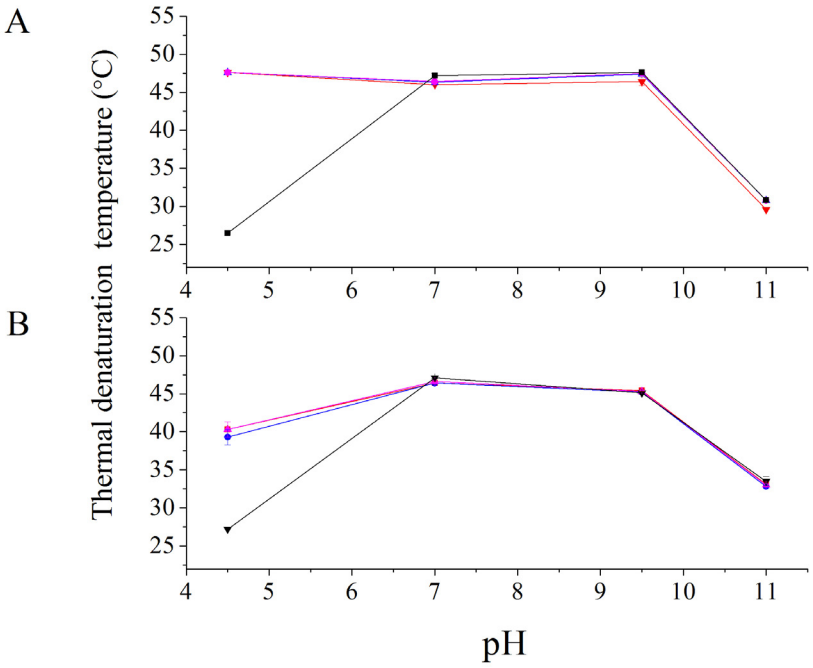


Figure 6.4 – The primary thermal denaturation temperatures of **(A)** pET 151/D-TOPO DHDPS Q90L/Y107W and **(B)** cleaved pET 151/D-TOPO DHDPS Q90L/Y107W in 100 mM phosphate buffer at 0 mM NaCl (red), 40 mM NaCl (blue), 100 mM NaCl (pink) and 500 mM NaCl (black) plotted as a function of pH. The mean thermal denaturation temperatures plotted were calculated from derivations of the fluorescence data. The data plotted are the mean values of six replicates and the error bars represent the SEM.

Under some conditions, the variants exhibited biphasic thermal denaturation data (see appendix 5, section A5.2). Such biphasic data have been attributed to the occurrence of two rate determining processes [7] as discussed in chapter 5, section 5.2.2.5. The presence of two oligomeric structures may have been responsible for two thermal denaturation temperatures with the larger arrangement undergoing thermal denaturation at a higher temperature than the smaller species [4, 5]. The possibility of abnormal protein folds which can also lead to such biphasic data, was investigated using CD spectroscopy and is described in section 6.2.2.5.

DSF demonstrated that the Q90L/Y107W double amino acid substitution has a significant effect on the thermal denaturation temperature of DHDPS compared to wild-type DHDPS and DHDPS Q90L. The temperature at which the double mutant thermally denatures is similar to DHDPS containing the Y107W parent mutation alone, supporting the assertion that the thermal stability is conferred primarily by the quaternary structure.

6.2.2.5 CIRCULAR DICHROISM SPECTROSCOPY

The secondary structures of the DHDPS Q90L/Y107W variants were assessed using CD spectroscopy and results were compared to those of wild-type DHDPS, pET 151/D-TOPO DHDPS Q90L and pET 151/D-TOPO DHDPS Y107W (see figure 6.5). The mean residue ellipticity at lower wavelengths is reduced compared to wild-type, and similar to the proteins containing the parent mutations. These results suggest that, as expected, the amino acid substitutions at positions 90 and 107 disrupt the secondary structure of DHDPS.

Thermal denaturation experiments were also performed by CD spectroscopy (as described in chapter 8, section 8.6.4) by monitoring the change in intensity at 220 nm. Both the DHDPS Q90L/Y107W variants demonstrated similar denaturation profiles to wild-type DHDPS and the DHDPS species containing the parent mutations (see appendix 5, section A5.3). Corroborating the results of the DSF analysis, the temperatures at which both double mutants thermally denatured were significantly lower than wild-type DHDPS and the proteins containing the parent mutation at position 90 (see table 6.2). The discrepancies between the results obtained by CD and DSF are small and consistent with previous findings [7-9] and those presented in chapter 5, sections 5.2.2.5 and 5.2.2.6.

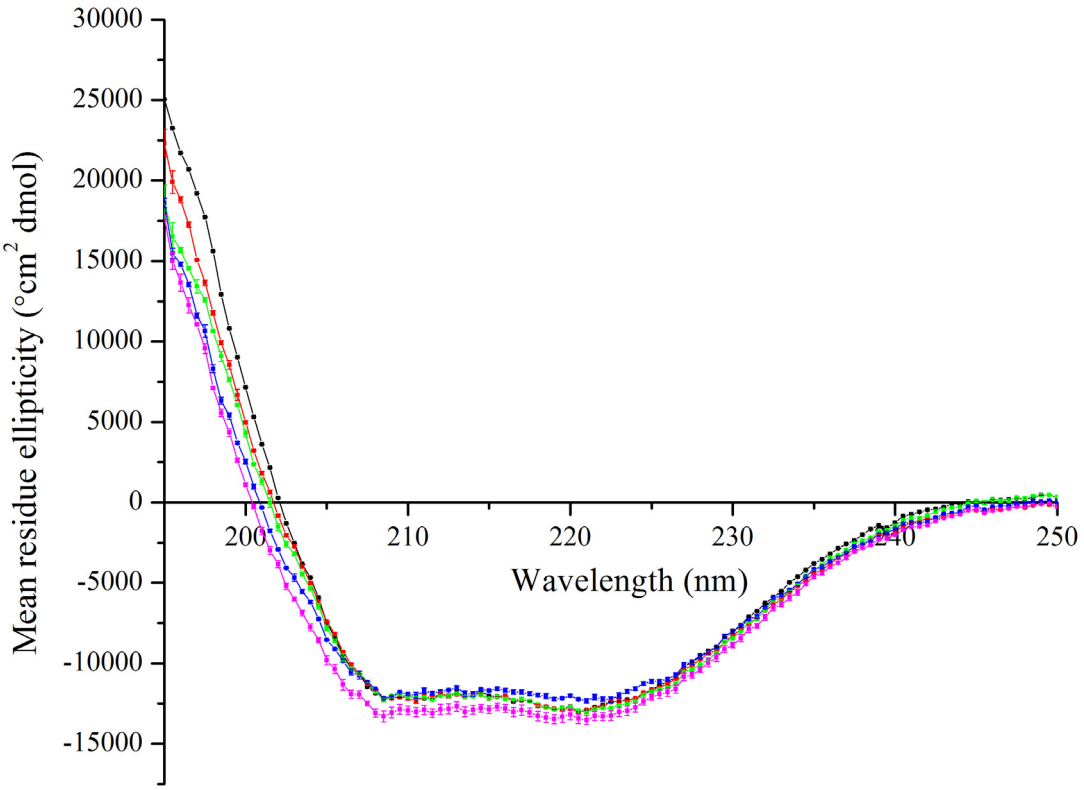


Figure 6.5 – Wavelength scans of wild-type DHDPS (black), pET 151/D-TOPO DHDPS Q90L (red), pET 151/D-TOPO DHDPS Y107W (green), pET 151/D-TOPO DHDPS Q90L/Y107W (blue) and cleaved pET 151/D-TOPO DHDPS Q90L/Y107W (pink) at 0.2 mg/mL in 20 mM phosphate buffer (pH 7) containing 40 mM. The plot shows mean residue ellipticity vs. wavelength ± SEM (n = 3).

Variant	**Thermal denaturation temperature as per CD (°C)**	**Thermal denaturation temperature as per DSF (°C)**
Wild type DHDPS	56.9 ± 0.15	60.90 ± 0.26
pET 151/D-TOPO DHDPS Y107W	47.5	46.1 ± 0.3
pET 151/D-TOPO DHDPS Q90L	59.9	61.7 ± 0.0
pET 151/D-TOPO DHDPS Q90L/Y107W	47.1	46.3 ± 0.1
Cleaved pET 151/D-TOPO DHDPS Q90L/Y107W	47.5	46.4 ± 0.0

Table 6.2 – The temperature at which wild-type DHDPS, pET 151/D-TOPO DHDPS Y107W, pET 151/D-TOPO DHDPS Y107W, pET 151/D-TOPO DHDPS Q90L/Y107W, and cleaved pET 151/D-TOPO DHDPS Q90L/Y107W melt as monitored by CD spectroscopy and DSF. There are no errors on the CD data as discussed in chapter 2, section 2.3.2.6 and the DSF data are that of the primary thermal denaturation temperature.

6.2.2.6 AMORPHOUS AGGREGATION

The amorphous aggregation propensities of the DHDPS Q90L/Y107W variants were measured under all chosen conditions (see section 2.2) by monitoring the light scattering at 340 nm. As the cleavage step resulted in a significant reduction in yield (see section 6.2.1) only three of the five salt conditions were screened for cleaved pET 151/D-TOPO DHDPS Q90L/Y107W. The results obtained for each variant were compared to the results for wild-type DHDPS and the DHDPS species containing the parent mutations. Each assay was carried out as described in chapter 8, section 8.6.5 and previous chapters.

The results of this assay indicate that the DHDPS Q90L/Y107W variants aggregate significantly faster than wild-type DHDPS (see figure 6.6). They also aggregate significantly faster than DHDPS Q90L. This result is similar to those obtained for the variants containing the amino acid substitution at position 107 alone, which demonstrated an increased speed of aggregation, attributed to the disruption of the quaternary structure and the reduction in thermal denaturation temperature (see chapter 5, section 5.2.2.7). The results presented in this section corroborate this hypothesis.

The effect of the full length pET 151/D-TOPO polyhistidine tag compared to the residues remaining following cleavage on the aggregation of DHDPS Q90L/Y107W follows the same pattern as all other variants. The cleavage of the polyhistidine motif reduced the maximum aggregation observed and caused a systematic decrease in $Agg_{½}$ at pH 4.5 with increasing salt concentration. In general, the effect of salt concentration on the polyhistidine tagged variants was similar to that observed for all other DHDPS variants. Salt had little effect on the $Agg_{½}$ except under the conditions mentioned above and only resulted in changes in the maximum aggregation at high concentrations (500 mM).

The effect of pH on the aggregation of the DHDPS Q90L/Y107W variants was significant. All variants exhibited the highest maximum light scattering at pH 4.5 followed by pH 7 and pH 9.5. This correlates with the results described in previous chapters. Very low readings were obtained for pH 2 and pH 11, except for in the presence of 500 mM NaCl, which appeared to cause an increase in aggregation. The aggregation half life at pH 4.5 was too rapid to quantify; however, the $Agg_{½}$ values for pH 7 and pH 9.5 were calculated and compared. The only significant difference occurred

between pH 7 and pH 9.5 at low salt concentrations. At all other salt concentrations the aggregation half life values were similar.

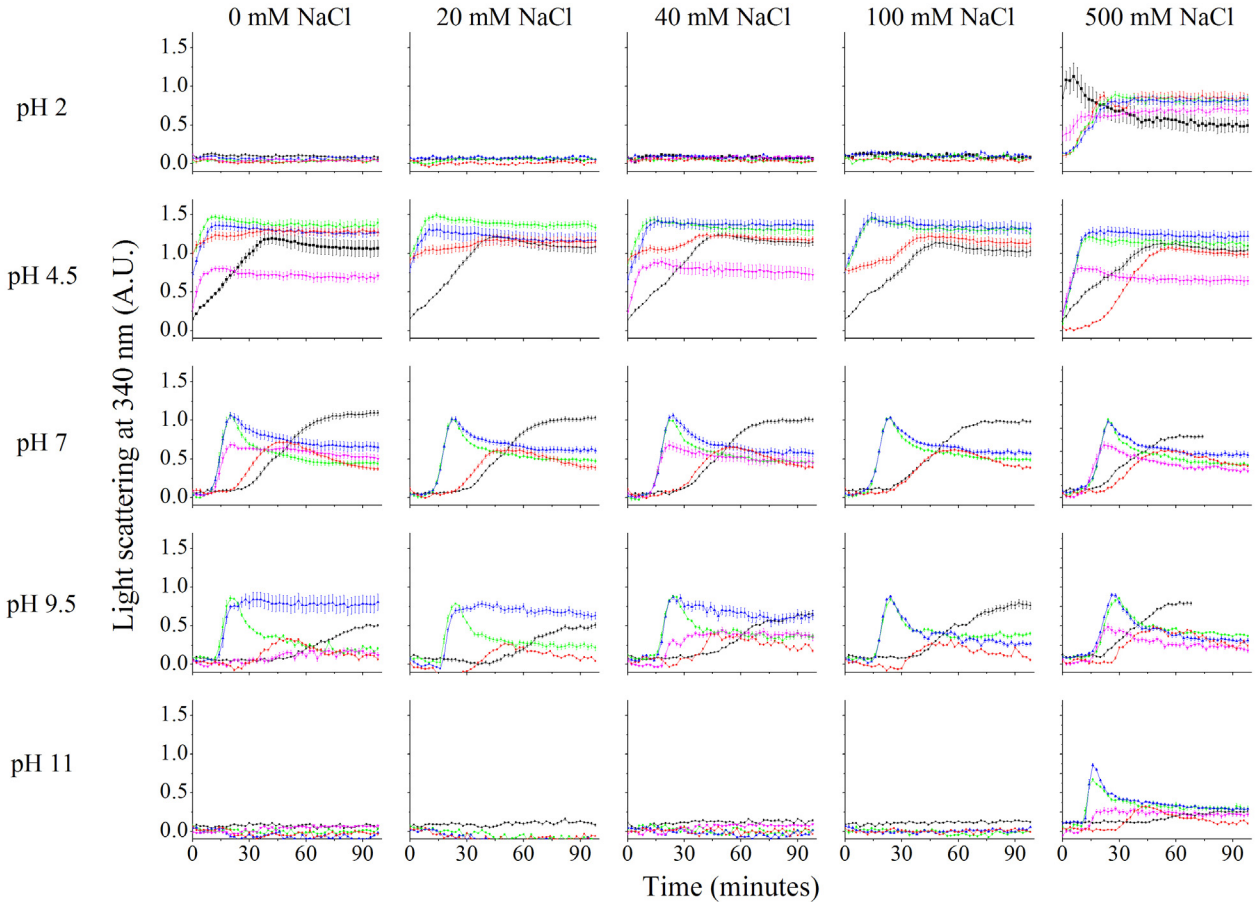


Figure 6.6 – The aggregation profiles of wild-type DHDPS (black), pET 151/D-TOPO DHDPS Q90L (red), pET 151/D-TOPO DHDPS Y107W (green), pET 151/D-TOPO DHDPS Q90L/Y107W (blue) and cleaved pET 151/D-TOPO DHDPS Q90L/Y107W (pink) at 0.5 mg/mL in 100 mM phosphate buffer. Where possible, a sigmoid (Boltzmann) function was used for the line of best fit (> 98% confidence interval) in order to calculate the aggregation half life ($Agg_{½}$). The $Agg_{½}$ was determined by identifying the midpoint of the transition as previously described [10]. In order to fit the sigmoid curve, only the data up until the linear phase was included. Once the peak absorbance was reached the light scattering was inconsistent due to the sedimentation of the protein aggregate. The maximum fluorescence was the highest reading over the initial 90 minutes of the assay. The data plotted are the mean values of six replicates and the error bars represent the SEM.

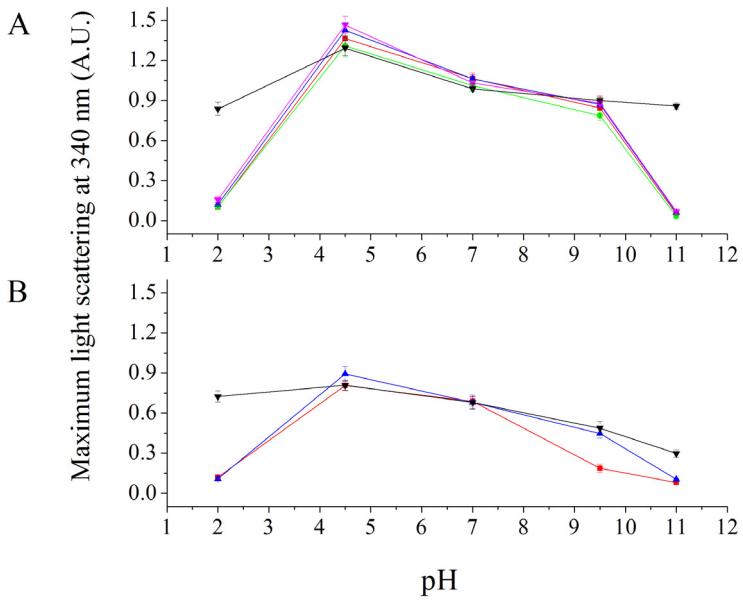


Figure 6.7 – The maximum absorbance of **(A)** pET 151/D-TOPO DHDPS Q90L/Y107W and **(B)** cleaved pET 151/D-TOPO DHDPS Q90L/Y107W at 0.5 mg/mL in 100 mM phosphate buffer 0 mM NaCl (red), 20 mM NaCl (green), 40 mM NaCl (blue), 100 mM NaCl (pink) and 500 mM NaCl (black). The data plotted are the maximum absorbance units reached over the initial 90 minutes of the assay and are plotted as a function of pH. The data plotted are the mean values of six replicates and the error bars represent the SEM.

The results of this assay indicated that the DHDPS Q90L/Y107W variants aggregate faster than wild-type DHDPS and DHDPS Q90L. This is an analogous result to those described for the DHDPS Y107W variants (chapter 5, section 5.2.2.7) and supports the theory that it is the quaternary structure that confers stability to DHDPS. From these results it appears that the Y107W amino acid substitution dictates the propensity to aggregate. The presence of the second amino acid substitution at position 90 does not appear to have destabilised the protein any further. However, the combinatorial effect of the amino acid substitutions on the propensity of DHDPS to form β-sheet specific aggregate were investigated and the results are described in the following section (section 6.2.2.7).

6.2.2.7 β-SHEET-SPECIFIC AGGREGATION

The propensities of the DHDPS Q90L/Y107W variants to form β-sheet rich aggregate were assessed across a range of conditions. Each assay was carried out as described chapter 8, section 8.6.6

and as described in previous chapters. The results were compared to those obtained for wild-type DHDPS and the DHDPS species containing the parent mutations.

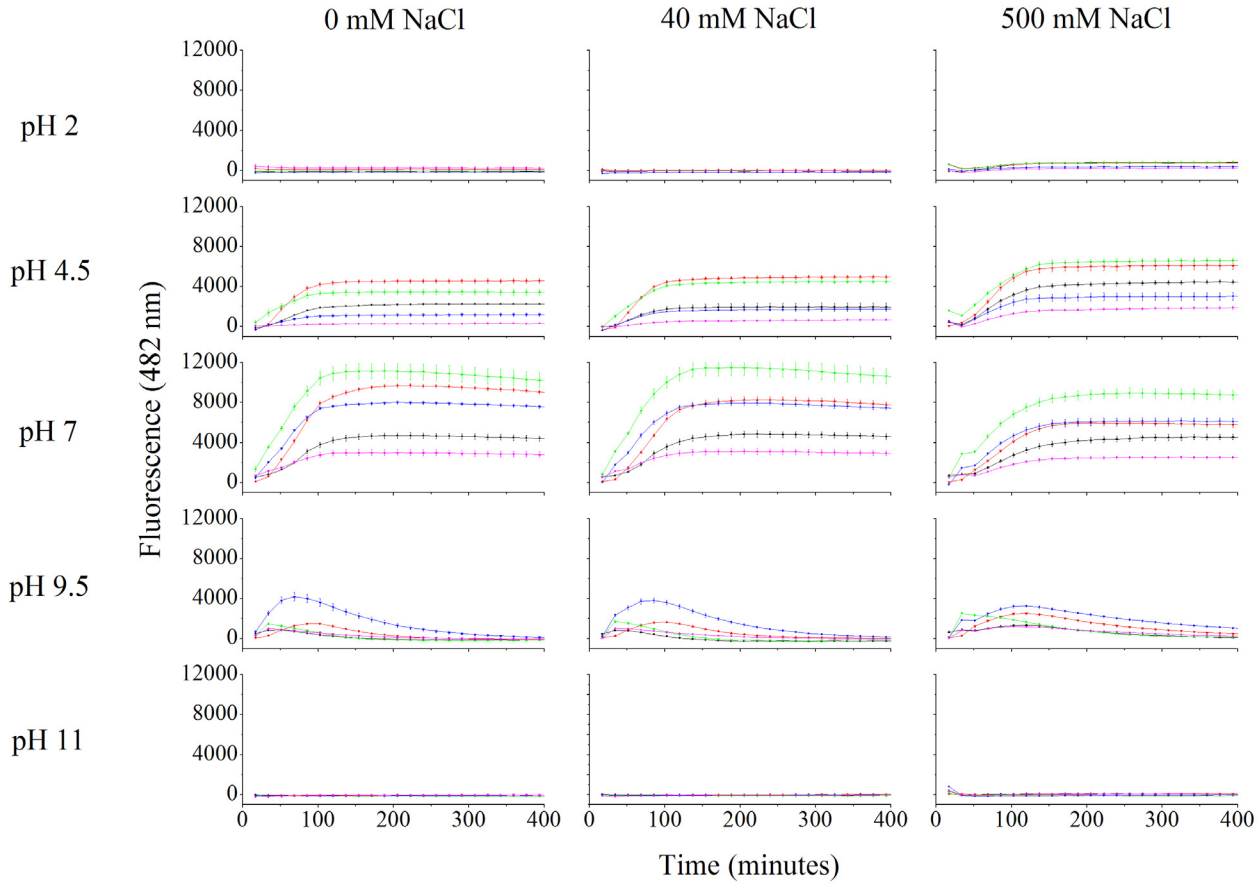


Figure 6.8 – The ThT fluorescence in the presence of wild-type DHDPS (black), pET 151/D-TOPO DHDPS Q90L (red), pET 151/D-TOPO DHDPS Y107W (green), pET 151/D-TOPO DHDPS Q90L/ Y107W (blue) and cleaved pET 151/D-TOPO DHDPS Q90L/Y107W (pink) in 100 mM phosphate buffer. The data plotted are the mean values of six replicates ± SEM. Where possible, a sigmoid (Boltzmann) function was used for the line of best fit (> 98% confidence interval) in order to calculate the β-sheet specific aggregation half life (β-$agg_{½}$). The β-$agg_{½}$ was determined by identifying the midpoint of the transition as previously described [10]. Note that only the data up until the linear phase was included in the analysis of the β-$agg_{½}$ [11]. Once the maximum fluorescence was reached the fluorescence readings were inconsistent due to sedimentation of the protein aggregate [11]. The maximum fluorescence was the highest reading over the initial 300 minutes of the assay.

The comparative analysis revealed that at pH 7, pET 151/D-TOPO DHDPS Q90L/Y107W is more prone to forming β-sheet rich aggregates that wild-type although less prone to inducing ThT fluorescence than either of the parent mutations. At pH 4.5 the fluorescence observed for this variant was consistently less than wild-type. Cleaved pET 151/D-TOPO DHDPS Q90L/Y107 exhibited less β-sheet rich aggregation propensity than wild type across all conditions. These results suggest that the combinatorial affects of the two amino acid substitutions resulted in a reduced tendency of the DHDPS to form β-sheet specific aggregates compared to the pET 151D-TOPO DHDPS variants containing the parental mutations. These results conflict with the predictions of Zyggregator which indicated an increase in propensity to form β-sheet specific aggregate upon double amino acid substitution.

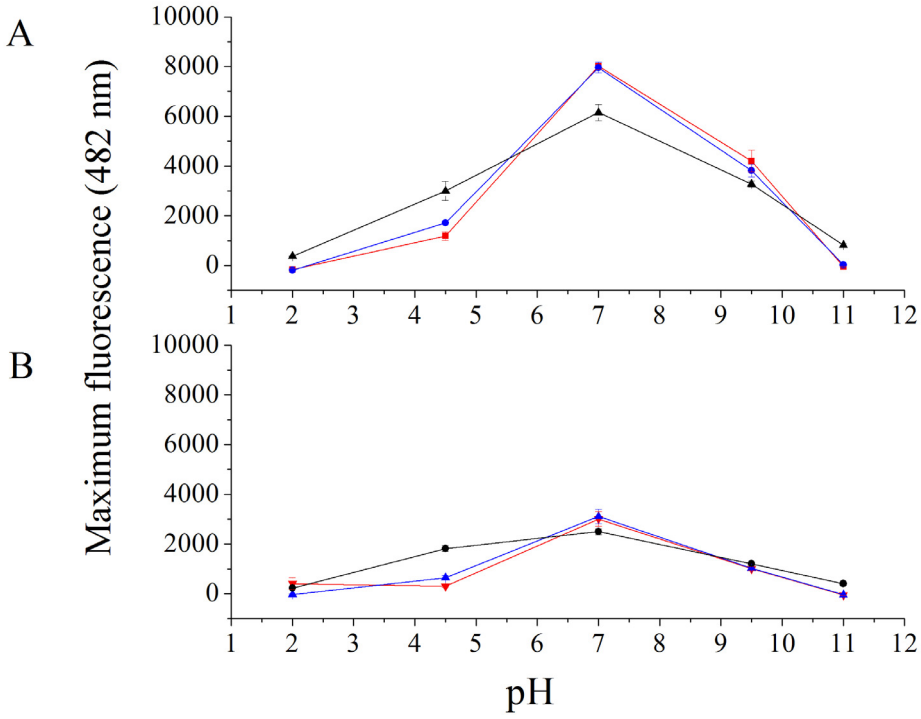


Figure 6.9 – The aggregation of **(A)** pET 151/D-TOPO DHDPS Q90L/Y107W and **(B)** cleaved pET 151/D-TOPO DHDPS in 0 mM NaCl (red), 40 mM NaCl (blue), and 500 mM NaCl (black). The data plotted are the mean maximum fluorescence (n=6) reached over 300 minutes and are plotted as a function of pH ± SEM.

From the data obtained from these assays, it appears that the two amino acid substitutions at positions 90 and 107 do not have an additive affect, exhibiting less fluorescence that the proteins

containing the parent mutations. To check for the presence of amyloid fibrils in the samples, TEM and X-ray fibre diffraction were employed. The results of these trials are described below.

6.2.2.8 CONFIRMATION OF AMYLOID FIBRILS

Attempts were made to confirm the presence of amyloid fibrils by TEM. Samples containing pET 151/D-TOPO DHDPS Q90L/Y107W were prepared under conditions which yielded the highest ThT fluorescence (pH 7, 0 mM NaCl and 40 mM NaCl) (as described in chapter 8, sections 8.8.1). Examination of the grids revealed no fibrillar structures (see figure 6.10) and attempts to create X-ray fibre diffraction stalks were unsuccessful. These two results suggest that the ThT results were not indicative of classical amyloid fibril formation and suggests that the additional presence of the Q90L amino acid substitution actually reduces any propensity of DHDPS Y107W to form fibrillar aggregates as observed for the parental Y107W variant (and as described in chapter 5, section 5.2.2.9).

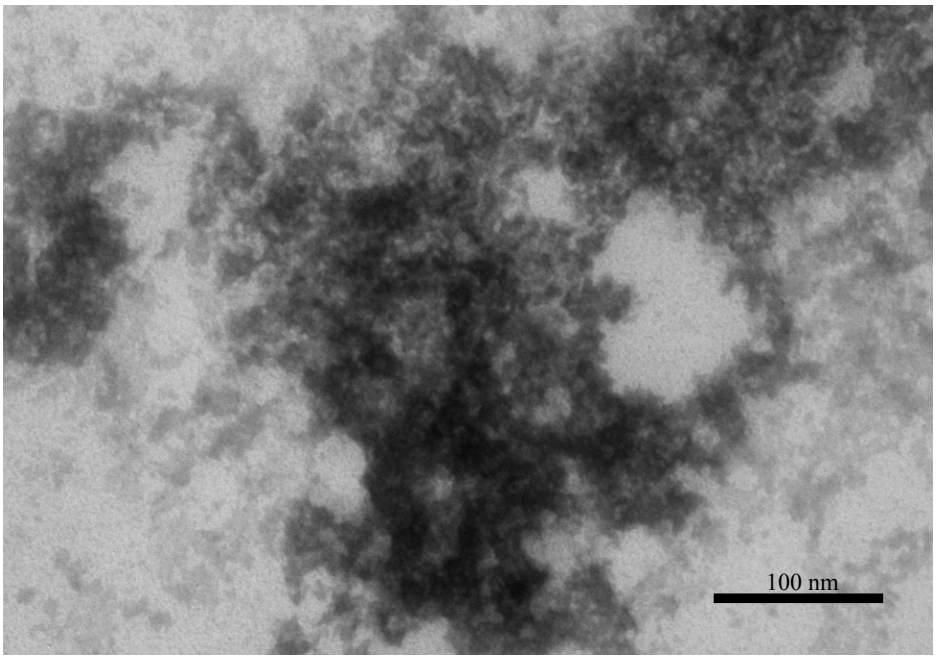


Figure 6.10 – TEM image of amorphous aggregate formed by pET 151/D-TOPO DHDPS Q90L/Y107W at 0.5 mg/mL in 100 mM phosphate buffer (pH 7), 40 mM NaCl at 60 ºC (120,000 x magnification). This image was typical of all 6 images obtained.

6.3 SUMMARY AND CONCLUSIONS

Upon comparison of the DHDPS variants containing the amino acid substitutions at both positions 90 and 107 with wild-type and the DHDPS variants containing the parent mutations (described in chapters 4 and 5), it was apparent that the primary determinant of many of the biophysical parameters was the quaternary structure of the enzyme. The second amino acid substitution at position 90 did not appear to significantly affect the kinetic parameters or the thermal denaturation temperature of DHDPS. This supports the results of the polyhistidine tagged versions of the parental mutants described in chapter 3.

The amorphous aggregation propensity of both the polyhistidine tagged and cleaved variants of DHDPS Q90L/Y107W was altered, with both proteins aggregating more rapidly than wild-type. In chapter 5, a similar increase in the speed of aggregation was attributed to the disruption of the quaternary structure and the resulting exposure of hydrophobic patches and the reduction in thermal denaturation temperature. Thus it appeared that the main determinant of the formation of aggregates was the amino acid substitution at position 107. The presence of leucine at position 90 in place of the glutamine does not appear to significantly affect the aggregation potential of DHDPS Q90L/Y107W, a result supported by the data presented in chapter 5 which found that Q90L in the absence of any polyhistidine motif, did not increase the aggregation propensity of DHDPS.

Both Q90L and Y107W individually caused a significant increase in the ThT induced fluorescence of DHDPS. Surprisingly, in combination, the fluorescence was reduced relative to both DHDPS Q90L and DHDPS Y107W suggesting that the propensity of the DHDPS Q90L/Y107W variants to form β-sheet-specific aggregates was lower than either DHDPS Q90L or DHDPS Y107W.

The results of the Zyggregator algorithm predicted the double amino acid substitution to result in a DHDPS species with an increased propensity to form amyloid fibrils. The results presented in this chapter suggest that Zyggregator is not accurate for DHDPS under these experimental conditions.

6.4 REFERENCES

1 Dobson, R. C. J., Valegard, K. and Gerrard, J. A. (2004) The crystal structure of three site-directed mutants of *Escherichia coli* dihydrodipicolinate synthase: Further evidence for a catalytic triad. Journal of Molecular Biology **338**, 329 - 339

2 Andrews, P. (1965) The gel filtration behaviour of proteins related to their molecular weights over a wide range. Biochemical Journal **96**, 595 - 606

3 Andrews, P. (1964) Estimation of the molecular weights of proteins by Sephadex gel filtration. Biochemical Journal **91**, 222 - 233

4 Pearce, F. G., Dobson, R. C., Weber, A., Lane, L. A., McCammon, M. G., Squire, M. A., Perugini, M. A., Jameson, G. B., Robinson, C. V. and Gerrard, J. A. (2008) Mutating the tight-dimer interface of dihydrodipicolinate synthase disrupts the enzyme quaternary structure: towards a monomeric enzyme. Biochemistry **In press**

5 Griffin, M. D., Dobson, R. C. J., Pearce, F. G., Antonio, L., Whitten, A. E., Liew, C. K., Mackay, J. P., Trewhella, J., Jameson, G. B., Perugini, M. A. and Gerrard, J. A. (2008) Evolution of quaternary structure in a homotetrameric enzyme. Journal of Molecular Biology **380**, 691 - 703

6 Vedadi, M., Niesen, F. H., Allali-Hassani, A., Fedorov, O. Y., Finerty, P. J. J., Wasney, G. A., Yeung, R., Arrowsmith, C., Ball, L. J., Berglund, H., Hui, R., Marsden, B. D., Norlund, P., Sundstrom, M., Weigelt, J. and Edwards, A. M. (2006) Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. Proceedings of the National Academy of Sciences USA **103**, 15835 - 15840

7 Epps, D. E., Sarver, R. W., Rogers, J. M., Herberg, J. T. and Tomich, P. K. (2001) The ligand affinity of proteins measured by isothermal denaturation kinetics. Analytical Biochemistry **292**, 40 - 50

8 Yeh, A. P., McMillan, A. and Stowell, M. H. B. (2006) Rapid and simple protein-stability screens: Applications to membrane proteins. Acta Crystallographica Section D: Biological Crystallography **D62**, 451 - 457

9 Leikina, E., Mertts, M. V., Kuznetsova, N. and Leikin, S. (2002) Type I collagen is thermally unstable at body temperature. Proceedings of the National Academy of Sciences USA **99**, 1314 - 1318

10 Sabate, R., Gallardo, M. and Estelrich, J. (2003) An autocatalytic reaction as a model for the kinetics of the aggregation of β-amyloid. Biopolymers (Peptide Science) **71**, 190 - 195

11 Niesen, F. H., Berglund, H. and Vedadi, M. (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. Nature Protocols **2**, 2212 - 2221

CHAPTER 7

SUMMARY AND CONCLUSIONS

The aim of this research was to advance the understanding of the factors that prevent $(\alpha/\beta)_8$ barrel proteins from aggregating. This was achieved through the introduction of various point mutations into *E. coli* DHDPS, a homotetrameric enzyme involved in the lysine biosynthetic pathway. $(\alpha/\beta)_8$ Barrels have never been shown to form amyloid fibrils. DHDPS was used as a model to test the hypothesis that the $(\alpha/\beta)_8$ barrel structure prevents proteins from unfolding and subsequently aggregating. The homotetrameric nature of DHDPS also enabled investigation into whether the quaternary structure played a protective role against aggregation and amyloid fibril formation.

The biophysical properties of DHDPS and its variants were assessed using a suite of methods. The secondary structures were assessed using CD spectroscopy and the tertiary fold was confirmed (where possible) by crystallographic data. Enzyme kinetics were used as a proxy indicator of tertiary and quaternary structure in cases where crystallographic analyses were not possible. The quaternary structures were established by analytical gel permeation liquid chromatography and, in some cases, AUC. The thermal denaturation temperatures of each variant were determined under a range of conditions using DSF and CD spectroscopy. The propensity of DHDPS and its variants to form amorphous aggregate was assessed though monitoring the increase in light scattering at 340 nm. The β-sheet-specific aggregation propensity was assessed by ThT fluorescence assays. Confirmation of the presence of linear aggregates was achieved through TEM. Wild-type DHDPS was characterised to provide "base-line" data to which all other variants were compared.

Protein concentration has been implicated in the propensity of proteins to aggregate and form amyloid fibrils. It has been established in humans that there is a strong correlation between expression levels and aggregation propensity, with those proteins most likely to form aggregates expressed at the lowest levels [1]. This suggests that in biological systems, the expression levels and propensity to aggregate are finely balanced. This balance may be disturbed in genetically modified organisms where concentrations of specific proteins are increased.

The modifications of proteins and the environments in which they exist can lead to unforeseen outcomes. The effect of the environment on the biophysical properties of DHDPS and its variants was investigated through the characterisation of the thermal denaturation temperatures and the propensity to form amorphous and β-sheet-specific aggregate across a range of pH and salt concentrations. What was apparent for all the mutants described was that the primary determinant of protein stability and misfolding was pH. The thermal denaturation temperatures for all DHDPS variants were highest at pH 7 and aggregation generally occurred the most rapidly at pH 4.5.

This chapter draws on all previous chapters and concludes that in addition to providing a functional role [2, 3] the evolution of the quaternary structure of DHDPS may prevent aggregation *in vitro*. Aggregation is also affected by the extra amino acids associated with polyhistidine tags. This has implications in biotechnology, which often employs such tags to facilitate purification of recombinant proteins.

7.1 ADDITION OF POLYHISTIDINE TAGS ALTERS THE BIOPHYSICAL PROPERTIES OF *E. COLI* DHDPS

Affinity tags are often utilised without considering the possible effects on the biophysical properties of the protein as a whole. Given that even single amino acid substitutions have been shown to dramatically alter the aggregation propensity of proteins [4-7], the impact of polyhistidine motifs (including the amino acid residues left behind following cleavage of the tag sequence), was likely to be significant.

The addition of the pET M11 and pET 151/D-TOPO polyhistidine tags to DHDPS did facilitate purification of most of the proteins of interest in this study; however, the kinetic parameters were different compared to the non-tagged enzymes. The secondary structures of the polyhistidine-tagged variants compared to wild-type were similar by CD spectroscopy. Crystallisation of the enzymes with full length tags was not achieved for any of the DHDPS variants, despite extensive screening. Polyhistidine tags reportedly do not usually prevent crystallisation [8]; however, their effect on crystal formation may be under-reported. In the case of DHDPS, it is possible that the presence of the polyhistidine tags shifted the equilibrium involved in crystal formation, resulting in proteins which were more prone to aggregate. This hypothesis was supported by empirical data which demonstrated

that the β-sheet-specific aggregation propensity of DHDPS was significantly increased for the full length pET M11 and pET 151/D-TOPO polyhistidine tags. ThT fluorescence is often used as a diagnostic tool for amyloid formation [9, 10]; however, TEM did not confirm these results, indicating that the ThT fluorescence was either a false positive [11] or indicative of an aggregate rich in β-sheet structures.

The cleaved pET 151/D-TOPO DHDPS was successfully crystallised and the X-ray crystal structure was solved by Pearce [12]. The pI values were not significantly altered and the protein exhibited no significant structural changes, leading to the hypotheses that either the increased mass had altered the crystallisation conditions (thus the equilibrium between the native and crystalline states) or alternatively, the disordered nature of the polyhistidine tags inhibited the formation of crystals. The additional amino acids associated with the polyhistidine tags did not significantly affect the quaternary structure or the thermal stability, suggesting that the increased aggregation propensity was due to the tag and was not due to the destabilisation of the protein.

The significant increase observed on the amorphous and β-sheet-specific aggregation propensity has implications for the use of polyhistidine tags. Although most biophysical parameters remained similar to that of the wild-type *in vitro*, the formation of β-sheet-rich aggregates may have impacts on how proteins with additional amino acids function and thus implications for how they may be applied in some biotechnologies.

7.2 ALGORITHMIC PREDICTION

At the outset of this work there were a limited number of prediction tools available for the assessment of protein stability and the effects of point mutations on aggregation propensity. Two algorithms were described in the introduction (chapter 1, section 1.6) and used to predict the aggregation propensity of the model protein, DHDPS. These algorithms consider similar biophysical parameters but utilise different modes of calculation. Interestingly, the results of Zyggregator and TANGO were dissimilar, identifying different regions of DHDPS as important in determining the overall β-sheet-specific aggregation propensity. The results of Zyggregator were chosen for testing due to the opportunity for collaboration with its developers with the view of providing empirical evidence for the further development of the algorithm.

The introduction of amino acid substitutions predicted to increase the β-sheet-specific aggregation propensity of DHDPS was carried out to investigate the hypothesis that the $(\alpha/\beta)_8$ barrel prevents proteins from forming amyloid fibrils. Three positions were chosen for mutation on the basis of the predictions of Zyggregator. These were the glutamine at position 90, the alanine at position 207 and the proline at position 270. These were altered to a leucine, valine and threonine respectively. Unfortunately, due to problems surrounding the expression and purification of P270T, characterisation of this variant had to be abandoned. It is possible that the presence of a threonine at position 270 increased the aggregation propensity of DHDPS *in vivo* and reduced yields. As P270 is highly conserved in DHDPS proteins from a wide range of organisms, it is likely that this residue plays a key role in the stabilisation or catalysis of DHDPS. There is some irony in the fact that potentially the most interesting DHDPS variant was also the most difficult to work with. With the advent of improved *in vivo* and *in vitro* expression systems, characterisation of this variant may eventually be possible.

The results of the biophysical analysis of the other two mutational variants (Q90L and A207V) demonstrated that the kinetic parameters and the structure of the proteins were similar to that of wild-type; however, Q90L did appear a little more disordered by CD spectroscopy, potentially explaining the difficulties encountered in obtaining crystals of this variant. DHDPS A207V, on the other hand, had a CD spectrum similar to wild-type and crystallised under the same conditions. The resulting crystal structure was solved and showed no significant changes to the tertiary fold. The quaternary structures of both Q90L and A207V were similar to wild-type and eluted as tetrameric DHDPS. The thermal denaturation temperatures and amorphous aggregation propensities were similar to wild-type; however, the β-sheet-specific aggregation propensity was reduced. These results indicate that the amino acid substitutions predicted by Zyggregator did not increase the amyloidogeneity of tetrameric DHDPS. As Zyggregator is designed to predict aggregation from an unfolded state and DHDPS remained folded in most of the conditions tested, the discrepancy between the theoretical and empirical data is not surprising.

Over the last few years, the prediction of aggregation from primary sequence alone has begun to be addressed. Algorithms have now been developed that take into account more than just simple sequence considerations [13-15]. These often require a PDB structure as input and based on structural features of the residues (such as the solvent accessibility, secondary structure and torsion angles), can predict the effect of mutations on the stability and aggregation propensity of the protein

as a whole. For this reason, they are likely to be more accurate for *in vitro* testing of aggregation propensity of naturally folded, globular proteins such as DHDPS. Zyggregator has recently been updated to take into account higher order structural considerations [16]. Several other such algorithms are described below.

RONN

RONN identifies disordered regions within protein structures and predicts their ability to affect the solubility of the protein as a whole. The algorithm is based on the dogma that some residues are typically associated with areas of disorder. These amino acid identifiers are used in conjunction with a pattern recognition algorithm that enables identification of potentially disordered regions within the protein [17].

CUPSAT

CUPSAT has been designed to identify point mutations that affect the stability of proteins. It uses the thermodynamic and structural properties (such as atom potentials and torsion angle distribution) of the amino acid chain to predict the difference in free energy of the unfolding of the wild-type protein compared to that containing the point mutant. It provides information regarding the structural features of specific residues and the stability of the protein upon point mutation of those residues to 19 alternative amino acids [14, 18].

I-Mutant2.0

I-Mutant2.0 predicts changes in protein stability upon single amino acid changes. The algorithm uses structural data and a neural network system to identify whether any one point mutation in question alters the relative stability of the protein as a whole [19]. It is able to utilise either structural or sequence data to carry out its predictions [13].

7.3 THE QUATERNARY STRUCTURE OF *E. COLI* DHDPS CONFERS STABILITY AGAINST AGGREGATION

The introduction of mutations that disrupt the quaternary structure of DHDPS have been shown to attenuate its catalytic activity in lysine biosynthesis [2, 3]. The buttressing of the dimers against one

another was hypothesised to optimise the dynamics of the protein, a factor increasingly recognised as critical for enzyme catalysis [20]. The quaternary structure of DHDPS was thus hypothesised to be an evolutionary adaptation that improved the activity and substrate specificity of an ancestral dimeric form of the protein [2] and was essential for its function in lysine biosynthesis. However, the recent description of the dimeric structure of DHDPS from MRSA has necessitated a re-evaluation of this hypothesis [21]. The MRSA DHDPS dimer is stabilised by the interaction of a large number of residues at the interface of the two monomers, which is proposed to be an alternative means of optimising dynamics across the interface. [2]. The existence of a fully functional dimeric DHDPS hints that the evolution of DHDPS in *E. coli*, and all other species [22-26], into a tetramer has other advantages.

Proteins have been described in which the quaternary structure appears to play a significant role in stabilisation and preventing aggregation. An example of this is TTR which, upon disruption of the tetramer, undergoes amyloid fibril formation more readily than the wild-type protein. Chen & Dokholyan hypothesised that proteins that form oligomeric complexes are subjected to higher selection pressure against forming aggregates compared to proteins that do not form oligomers [27]. During the course of this thesis, an alternative hypothesis was proposed. This was that oligomeric structures evolved to prevent proteins from misfolding and amyloid fibril formation. To investigate this, the role of the homotetrameric nature of DHDPS in the prevention of aggregation and amyloid formation was investigated.

Two point mutants (Y107W and L197Y) had previously been identified as having aberrant quaternary structures [2, 3]. The introduction of a tyrosine at position 197 in place of a leucine results in a dimeric DHDPS [2]. The substitution of the tyrosine at position 107 with a tryptophan results in a DHDPS that exits in a monomer/tetramer equilibrium [3]. These proteins provided an ideal model to probe the relationship between the tetrameric nature of DHDPS and the stability and aggregation propensity. Many of the biophysical properties of Y107W and L197Y had already been described; however, despite anecdotal evidence that these variants were more prone to aggregation, their aggregation propensities had not been quantified. As Y107 is involved in the active site, a third mutant was characterised (Y107F). This variant had previously been described [28] and, because it does not alter the quaternary structure, was used as a control for mutation at position 107.

In order for direct comparisons to be made to the other variants, properties such as the kinetics, the thermal denaturation temperature, the quaternary structure and the CD spectra were investigated. Generally, the results obtained during the course of this research were similar to that reported in the literature [2, 3, 28]. Analysis of the quaternary structure revealed that the Y107W and L197Y variants were not tetrameric. By analytical gel permeation liquid chromatography, L197Y was dimeric, Y107W was in a monomer-tetramer equilibrium and DHDPS Y107F was a tetramer in solution. As expected, the disruption of the quaternary structure significantly decreased the catalytic ability of the enzyme, as did the substitution of the tyrosine at position 107 with a phenylalanine, reflective of its involvement in the active site.

The comparison of the published crystal structures with that of the wild-type enzyme revealed no significant changes to the tertiary structures; however, the CD spectroscopic analysis suggested that there was some disorder in the variants. The disrupted quaternary structure had previously been shown to reduce the thermal stability of DHDPS [2, 3], an observation confirmed by CD and DSF. The reduced thermal denaturation temperatures were associated with rapid formation of amorphous aggregate. The propensity of both Y107W and L197Y to form β-sheet-specific aggregate was also increased, suggesting that the hypothesis that the quaternary structure prevents amorphous and β-sheet-specific aggregation was correct.

The hydrophobic residues involved in the dimer-dimer interfaces of DHDPS are exposed upon disruption of the quaternary structure, likely facilitating the aggregation processes. Whether the tetrameric nature of DHDPS evolved to mask the interface residues or the presence of the hydrophobic residues at the interface evolved to facilitate subunit interaction cannot be established. The increased disorder of the quaternary structure mutants observed by CD spectroscopy was possibly due to the loss of the stabilising influence of the quaternary structure. Such increased movement may elevate the likelihood of unfolding, a process thought to be critical to aggregation and amyloid formation of globular proteins [29]. What can be established is that, whether aggregation is driven by the association of residues usually buried in the interfaces between monomers, or through destabilisation of the monomeric subunits resulting in unfolding, the tetrameric structure of DHDPS (as with the case of TTR) prevents misfolding.

7.4 CHARACTERISTATION OF A DOUBLE MUTANT REVEALS THAT QUATERNARY STRUCTURE IS THE MAJOR DETERMINANT OF STABILITY.

Zyggregator predicts aggregation from an unfolded, monomeric state. For this reason, a double mutant was created with an aberrant quaternary structure due to the presence of Y107W. It also contained the amino acid substitution, Q90L. It was hypothesised that the accuracy of Zyggregator would improve with respect to the predictions for Q90L in a non-homotetrameric DHDPS species.

Characterisation of the Q90L/Y107W variants revealed that the kinetic parameters were similar to those of DHDPS Y107W. This was foreseen as the quaternary structure had already been established as the major determinate of catalytic efficiency. Analysis revealed that the tetrameric nature of the enzyme was disrupted and that the thermal denaturation temperatures were similar to DHDPS Y107W. The presence of the amino acid substitution at position 90 appeared to have no further destabilising affect. In fact, the presence of the amino acid mutation at position 90 appears to have reduced amorphous and β-sheet-specific aggregation propensity compared to DHDPS Y107W, suggesting that the presence of a leucine at position 90 stabilised DHDPS. This postulate is supported by the evidence presented for DHDPS Q90L.

7.5 THE COMBINATORIAL AFFECTS OF MUTATIONS, POLYHISTIDINE TAGS AND ENVIRONMENT MAY HAVE UNFORESEEN EFFECTS

Polyhistidine tags were used to facilitate the purification of the variants of DHDPS. As observed for the wild-type enzyme, the kinetic parameters were altered; generally decreasing the k_{cat}^{app} and increasing the K_M^{app} with respect to pyruvate compared to wild-type DHDPS and the non-tagged variants. In addition to this, the ability of the polyhistidine-tagged variants to crystallise was compromised, indicating a shift in the equilibria involved. The amino acid substitutions at positions 90, 107, 197 and 207 exhibited different behaviours with regard to their propensity to aggregate dependent on the sequence of the polyhistidine motif. For example, the polyhistidine motif appeared to alter the monomer-tetramer equilibrium exhibited by DHDPS Y107W, resulting in a primarily

monomeric species by gel filtration. This suggested that the amino acids associated with the polyhistidine motifs were affecting the equilibrium between the tetramer and the monomer of DHDPS Y107W but only in the gel matrix of the chromatography column. The pET M11 polyhistidine tag increased the β-sheet-specific aggregation of DHDPS Q90L and decreased the fluorescence of DHDPS A207V. The pET 151-D/TOPO polyhistidine tag generally increased ThT fluorescence for both Q90L and A207V. While this increase may have indicated the presence of β-sheet structures in solution, the formation of amyloid fibrils could not be confirmed by TEM. The ThT fluorescence was also high for the polyhistidine-tagged Y107W variants and elevated for the pET 151/D-TOPO DHDPS L197Y and its cleaved variant. Using TEM, linear aggregates, consistent with the appearance of amyloid fibrils, were observed. Although cross β-structures were not confirmed by X-ray fibre diffraction, these results suggest that in the presence of the pET M11 polyhistidine tag, the quaternary structure of DHDPS may prevent amyloid formation.

The mechanisms underlying folding and misfolding of proteins are not well understood. From the results of this research it is apparent that the destabilisation and subsequent aggregation of proteins is not predictable. The fact that the polyhistidine tags increased the predisposition of some variants to form amorphous and β-sheet-specific aggregates is of particular interest. There have been previous reports of fusion proteins with increased propensities to aggregate. For example, the addition of the "degradon" peptide to GFP resulted in a fusion protein which formed intracellular deposits in several species. The presence of these intracellular deposits had detrimental effects *in vivo* and were thus described as cytotoxic [31].

These findings have implications for the expression, purification and application of fusion proteins. The increase in propensity of any protein to form β-sheet-specific aggregate or amyloid could increase the risk of environmentally acquired amyloidoses. It has been shown that unrelated amyloid-forming protein species can accelerate the progression of amyloid diseases *in vivo* [32]. This has serious implications not only regarding the addition of polyhistidine tags but, potentially, the addition of any motif to proteins.

A pertinent example is the commercial maize line LY038. As discussed in chapter 1, this contains the *dapA* gene obtained from *C. glutamicum*. It is also fused to an import sequence. Not only could the expression of the DHDPS from *C. glutamicum* (cDHDPS) in a novel physiological context and

concentration potentially lead to undesirable misfolding events, but the extra amino acids associated with the import sequence may also change the propensity of the enzyme to misfold.

7.6 FURTHER WORK

The optimisation of the expression of P270T would enable further characterisation of its properties. Due to its highly conserved nature, its presence near the active site and its involvement at the tight dimer interface it is likely to be an interesting variant of DHDPS.

Confirmation of the formation of a cross-β structure by DHDPS Y107W by X-ray fibre diffraction would conclusively establish that $(\alpha/\beta)_8$ barrels can form amyloid fibrils. This research was designed to investigate the aggregation landscape of DHDPS in general. More extensive screening of Y107W under conditions that generally promote the formation of amyloid (*e.g.* molecular crowding, increased concentration, and the presence of chaotropic agents) may provide additional information as to what other factors can induce aggregation and fibril formation of DHDPS.

Testing of algorithms such as Zyggregator that consider structures beyond just the primary sequence would prove interesting, particularly in the case of $(\alpha/\beta)_8$ barrels, and further investigation into the influence of the quaternary structure on preventing amyloid fibril formation in other systems would be a fascinating avenue of research.

The characterisation of cDHDPS with and without the import sequence may yield interesting results with applications to biosafety.

7.7 REFERENCES

1 Tartaglia, G. G., Pechmann, S., Dobson, C. M. and Vendruscolo, M. (2007) Life on the edge: a link between gene expression levels and aggregation rates in humans. Trends in Biochemical Sciences **32**, 204 - 206

2 Griffin, M. D., Dobson, R. C. J., Pearce, F. G., Antonio, L., Whitten, A. E., Liew, C. K., Mackay, J. P., Trewhella, J., Jameson, G. B., Perugini, M. A. and Gerrard, J. A. (2008) Evolution of quaternary structure in a homotetrameric enzyme. Journal of Molecular Biology **380**, 691 - 703

3 Pearce, F. G., Dobson, R. C., Weber, A., Lane, L. A., McCammon, M. G., Squire, M. A., Perugini, M. A., Jameson, G. B., Robinson, C. V. and Gerrard, J. A. (2008) Mutating the tight-dimer interface of dihydrodipicolinate synthase disrupts the enzyme quaternary structure: towards a monomeric enzyme. Biochemistry **In press**

4 Mishima, T., Ohkuri, T., Monji, A., Imoto, T. and Ueda, T. (2006) Amyloid formation in denatured single-mutant lysozymes where residual structures are modulated. Protein Science **15**, 2448 - 2452

5 Steward, R. E., Armen, R. S. and Daggett, V. (2008) Different disease-causing mutations in transthyretin trigger the same conformational conversion. Protein Engineering, Design and Selection **21**, 187 - 195

6 Chiti, F., Calamai, M., Taddei, F., Stefani, M., Ramponi, G. and Dobson, C. M. (2002) Studies of the aggregation of mutant proteins *in vitro* provide insights into the genetics of amyloid diseases. Proceedings of the National Academy of Sciences USA **99**, 16419 - 16426

7 Fernandez-Escamilla, A.-M., Rousseau, F., Schymkowitz, J. and Serrano, L. (2004) Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. Nature Biotechnology **22**, 1302-1306

8 Smyth, D. R., Mrokiewicz, M. K., McGrath, W. J., Listwan, P. and Kobe, B. (2003) Crystal structures of fusion proteins with large-affinity tags. Protein Science **12**, 1313 - 1322

9 Eisert, R., Felau, L. and Brown, L. R. (2006) Methods for enhancing the accuracy and reproducibility of congo red and thioflavin T assays. Analytical Biochemistry **353**, 144 - 146

10 Voropai, E. S., Samtsov, M. P., Kaplevskii, K. N., Maskevich, A. A., Stepuro, V. I., Povarova, O. I., Kuznetsova, N., Turoverov, K. K., Fink, A. L. and Uversky, V. N. (2003) Spectral properties of thioflavin T and its complexes with amyloid fibrils. Journal of Applied Spectroscopy **70**, 868 - 874

11 Groenning, M., Norrman, M., Flink, J. M., van der Weert, M., Butrinsky, J. T., Schluckebier, G. and Frokjaer, S. (2007) Binding of thioflavin T in insulin amyloid fibrils. Journal of Structural Biology **159**, 483 - 497

12 Pearce, F. G. (2008) Crystal structure of cleaved pET 151/D-TOPO polyhistidine tagged *E. coli* DHDPS. Personal communication, Christchurch

13 Capriotti, E., Fariselli, P. and Casadio, R. (2005) I-Mutant2.0: Predicting stability changes upon mutation form the protein sequence or structure. Nucleic Acids Research **33**, W306 - W310

14 Parthiban, V., Gromiha, M. M. and Schomburg, D. (2006) CUPSAT: prediction of protein stability upon point mutation. Nucleic Acids Research **34**, W239 - W242

15 Stitziel, N. O., Binkowski, A., Tseng, Y. Y., Kasif, S. and Liang, J. (2004) topoSNP: a topographic database of non-synonymous single nucleotide polymorphisms with and without known disease association. Nucleic Acids Research **32**, D520 - D522

16 Tartaglia, G. G. and Vendruscolo, M. (2008) The Zyggregator method for predicting protein aggregation propensities. Chemical Society Reviews **37**, 1395 - 1401

17 Yang, Z. R., Thomson, R., McNeil, P. and Esnouf, R. M. (2005) RONN: The bio-basis function neural network technique applied to the detection of natively disordered regions in proteins. Bioinformatics **21**, 3369 - 3376

18 Parthiban, V., Gromiha, M. M., Hoppe, C. and Schomburg, D. (2007) Structural analysis and prediction of protein mutant stability using distance and torsion potentials: Role of secondary structure and solvent accessibility. Protein: Structure, Function and Bioinformatics **66**, 41 - 52

19 Capriotti, E., Fariselli, P. and Casadio, R. (2004) A neural-network-based method for predicting protein stability changes upon single point mutations. Bioinformatics **20**, i63 - i68

20 Hammes-Schiffer, S. and Benkovic, S. J. (2006) Relating protein motion to catalysis. Annual Review of Biochemistry **75**, 519 - 541

21 Burgess, B., Dobson, R. C. J., Bailey, M. F., Atkinson, S. C., Griffin, M. D. W., Jameson, G. B., Parker, M. W., Gerrard, J. A. and Perugini, M. A. (2008) Structure and evolution of a novel dimeric enzyme from a clinically-important bacterial pathogen. Journal of Biological Chemistry **283**, 27598 - 27603

22 Pearce, F. G., Perugini, M. A., McKercher, H. J. and Gerrard, J. A. (2006) Dihydrodipicolinate synthase from *Thermotoga maritima*. Biochemical Journal **400**, 359 - 366

23 Kefala, G., Evans, G. L., Griffin, M. D., Devenish, S., Pearce, F. G., Perugini, M. A., Gerrard, J. A., Weiss, M. S. and Dobson, R. C. (2008) Crystal structure and kinetic study of dihydrodipicolinate synthase from *Mycobacterium tuberculosis*. Biochemical Journal **411**, 351 - 360

24 Blickling, S., Beisel, H.-G., Bozic, D., Knablein, J., Laber, B. and Huber, R. (1997) Structure of dihydrodipicolinate synthase of *Nicotiana sylvestris* reveals novel quaternary structure. Journal of Molecular Biology **274**, 608 - 621

25 Kumarevel, T. S., Karthe, P., Kuramitsu, S. and Yokoyama, S. Crystal structure of dihydrodipicolinate synthase from *Aquifex aeolicus*. Publication in PDB pending

26 Satyanarayana, L., eswaramoorthy, S., Sauder, J. M., Burley, S. K. and Swaminathan, S. Crystal structure of dihydrodipiolinate synthase from *Oceanbacillus iheyensis* at 1.9 Å resolution. Publication in PDB pending

27 Chen, Y. and Dokholyan, N. V. (2008) Natural selection against protein aggregation on self-interacting and essential proteins in yeast, fly and worm. Molecular Biology and Evolution **25**, 530 - 1533

28 Dobson, R. C. J., Valegard, K. and Gerrard, J. A. (2004) The crystal structure of three site-directed mutants of *Escherichia coli* dihydrodipicolinate synthase: Further evidence for a catalytic triad. Journal of Molecular Biology **338**, 329 - 339

29 Chiti, F. and Dobson, C. M. (2006) Protein misfolding, functional amyloid and human disease. Annual Review of Biochemistry **75**, 333 - 366

30 Mirwaldt, C., Korndorfer, I. and Huber, R. (1995) The crystal structure of dihydrodipicolinate synthase from *Escherichia coli* at 2.5 Å resolution. Journal of Molecular Biology **246**, 227 - 239

31 Link, C. D., Fonte, V., Hiester, B., Yerg, J., Ferguson, J., Csontos, S., Silverman, M. A. and Stein, G. H. (2006) Conversion of green fluorescent protein into a toxic aggregation-prone protein by the C-terminal addition of a short peptide. The Journal of Biological Chemistry **281**, 1808 - 1816

32 Lundmark, K., Westermark, G. T., Olsen, A. and Westermark, P. (2005) Protein fibrils in nature can enhance amyloid protein A amyloidosis in mice: Cross-seeding as a disease mechanism. Proceedings of the National Academy of Sciences USA **102**, 6098 - 6102

CHAPTER 8

EXPERIMENTAL MATERIALS AND METHODS

8.1 MATERIALS

Chemicals were purchased from Sigma-Aldrich Co. (Auckland, New Zealand) unless otherwise specified. Media for bacterial cultures were purchased from Invitrogen (Auckland, New Zealand). SDS PAGE gels were purchased from Invitrogen (Auckland, New Zealand. Restriction enzymes were purchased from Roche (Auckland, New Zealand). The BioRad protein assay kit was purchased from BioRad Laboratories (Auckland, New Zealand). DNA HyperLadder I was purchased from Total Lab Systems (Auckland, New Zealand). Column chromatography media were purchased from GE Healthcare (Auckland, New Zealand) or BioRad Laboratories (Auckland, New Zealand) as pre-packed columns. (*S*)-ASA was synthesised using a technique described in Roberts, 2003 [1] (adapted from Coulter, 1996 [2]) and provided by members of the laboratory (Dr Sean Devenish and Dr Andrew Muscroft-Taylor).

8.2 MICROBIOLOGICAL AND MOLECULAR METHODS

8.2.1 BACTERIAL STRAINS

Two bacterial strains were used in this research. *E. coli* XL-1 Blue (genotype *rec*A1 *end*A1 *gyr*A96 *thi*-1 *hsd*R17 *sup*E44 *rel*A1 *lac* [F' *pro*AB *lac*1 ZΔM15 Tn10 (*tet*^{r})]) was obtained from Stratagene as part of a QuikChange site-directed mutagenesis kit [3]. *E. coli* CodonPlus BL21 DE3 (RP) (genotype B F– *omp*T *hsd*S(r_B– m_B–) dcm+ *tet*^{r} *gal* λ (DE3) *end*A Hte [*arg*U *pro*L Cam^r]) was obtained from Stratagene.

8.2.2 PLASMIDS

The genes *dap*A and *dap*B encode dihydrodipicolinate synthase and dihydrodipicolonate reductase respectively. These genes had previously been cloned into pBluescript KS+ vectors by others [4, 5] and have subsequently been cloned into pET vectors in this laboratory [6].

Plasmid	Genotype	Source
pJG001	::*dap*A, *amp*^{r}	Lab stocks [4]
pET M11 dapA	:: his *dap*A, *kan*^{r}	Created from pET M11 [7]
pET 151/D-TOPO dapA	:: his *dap*A, *kan*^{r}	Created from pET 151/D-TOPO [8]
pJG001 Q90L	::*dap*A-gln90leu, *amp*^{r}	Created from pJG001
pET M11 Q90L	:: his *dap*A-gln90leu, *kan*^{r}	Created from pET M11 dapA
pET 151/D-TOPO Q90L	:: his *dap*A-gln90leu, *kan*^{r}	Created from pET 151/D-TOPO dapA
pJG001 Y107W	::*dap*A-try107trp, *amp*^{r}	Created from pJG001 by Pearce[9]
pET M11 Y107W	:: his *dap*A-try107trp, *kan*^{r}	Created from pET M11 dapA
pET 151/D-TOPO Y107W	:: his *dap*A-try107trp, *kan*^{r}	Created from pET 151/D-TOPO dapA
pJG001 Y107F	::*dap*A-try107phe, *amp*^{r}	Created from pJG001 by Dobson [10]
pJG001 L197Y	::*dap*A-leu197try, *amp*^{r}	Created from pJG001 by Griffin [11]
pET M11 L197Y	:: his *dap*A-leu197try, *kan*^{r}	Created from pET M11 dapA
pET 151/D-TOPO L197Y	:: his *dap*A-leu197try, *kan*^{r}	Created from pET 151/D-TOPO dapA
pJG001 A207V	::*dap*A-ala207val, *amp*^{r}	Created from pJG001
pET M11 A207V	:: his *dap*A-ala207val, *kan*^{r}	Created from pET M11 dapA
pET 151/D-TOPO A207V	:: his *dap*A-ala207val, *kan*^{r}	Created from pET 151/D-TOPO dapA
pET 151/D-TOPO Q90L/Y107W	:: his *dap*A-gln90leu, try107trp, *kan*^{r}	Created from pET 151/D-TOPO Y107W
pET M11 dapB	:: his *dap*B, *kan*^{r}	Created from pET M11 [6]
pUC18	:: *lac*Z, bla (*amp*^{r}), *rep* (pMB1)	Invitrogen [12]
pWhitescript	:: *lac*Z, *amp*^{r}	Stratagene [3]
pRK793	:: TEV protease – ser219val, *amp*^{r}	Addgene [13]

Table 8.1 – The plasmids utilised during the course of this research. All plasmids were created during the course of this study unless otherwise indicated.

8.2.3 BACTERIAL CULTURES

All media and equipment used for bacterial culturing were sterilised by autoclaving at 121 ºC for 20 minutes. Heat labile antibiotics and nutritional supplements were sterilised by filtration through 0.22 µm Millex GP filter units (Millipore, U.S.A.). Solutions were prepared using sterile, distilled water. Sterile technique was used during all manipulations of bacterial cultures and supplements.

8.2.4 MEDIA

Luria-Bertani medium (LB)

LB base (Gibco BRL) was supplied in a powdered form. As per the instructions, 20 g of LB base was added per 1 L sterile, distilled water and the pH was adjusted to 7.0 through addition of NaOH. Following this, the medium was autoclaved.

NZY+ broth

This broth was made up by adding 10 g NZ amine, 5 g yeast extract, and 5 g NaCl to 1 L sterile distilled water [14]. The pH was adjusted to 7.0 through addition of NaOH. Following this, the medium was autoclaved. Prior to use, 12.5 mL sterile 1 M $MgCl_2$, 12.5 mL sterile $MgSO_4$, and 20 mL sterile 20 % (w/v) glucose were added to the media.

Super optimal broth (SOB)

This broth was prepared in sterile, distilled water to a final concentrations of 2 % (w/v) bactotryptone, 0.5 % (w/v) yeast extract, 10 mM NaCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, and 0.36 % (w/v) glucose [14]. The pH was adjusted to 7.0 through addition of NaOH, then autoclaved.

Super optimal broth for catabolic repression (SOC)

SOC broth was prepared by supplementing SOB broth with 20 mM sterile glucose [14].

Auto-induction media (ZYM5052) [15]

This broth was prepared in 960 mL sterile, distilled water to a final concentration of 1 % (w/v) tryptone, 0.5 % (w/v) yeast extract. After autoclaving, 20 mL of sterile 50 x 5052 stock, 20 mL of 50 x M stock, 1 mL of 1 M $MgSO_4$ and 200 µL of 5000 x trace metals stock were added. The 50 x 5052 stock contained 25 % glycerol, 2.5 % glucose and 10 % lactose. The 50 x M stock contained 1.25 M Na_2HPO_4, 1.25 M KH_2PO_4, 2.5 M NH_4Cl, and 0.25 M Na_2SO_4. The 5000 x trace metals stock solution contained 20 mM $CaCl_2$, 10 mM $MnCl_2$, 10 mM $ZnSO_4.7H_2O$, 2 mM $CoCl_2.6H_2O$, 2 mM $CuCl_2.2H_2O$, 2 mM $NiCl_2.6H_2O$, 2 mM $Na_2MoO_4.2H_2O$, 2 mM $Na_2SeO_3.5H_2O$, 2 mM H_3BO_3, 50 mM $FeCl_3.6H_2O$.

8.2.5 ANTIBIOTICS AND NUTRITIONAL SUPPLEMENTS

Antibiotic and nutritional supplement stocks were sterilised by filtration and stored at -20 ºC. The concentrations of the stocks and the working concentrations are shown in Table 8.2.

Supplement	Stock concentration	Working concentration
Ampicillin	100 mg/mL	100 µg/mL
Tetracycline	20 mg/mL	20 µg/mL
Kanamycin	30 mg/mL	30 µg/mL
IPTG	1 M	1 mM

Table 8.2 – Antibiotic stock and working concentrations.

8.2.6 PLATE PREPARATION

LB medium for plates contained 15 g/L agar and was sterilised by autoclave. The medium, once cooled to ~50 ºC, was supplemented with the appropriate antibiotics and poured into sterile Petri dishes under a laminar flow hood. The plates were left to set for ~1 hour, sealed and then stored at 4 ºC. Blue/white selection plates were prepared by spreading 8 µL of 10 mg/mL IPTG and 40 µL of 20 mg/mL X-gal (in DMSO) onto the surface of LB plates. In order to remove residual DMSO, the plates were allowed to dry for a minimum of 2 hours in a laminar flow hood [14].

8.2.7 BACTERIAL CULTURING

Agar plates were inoculated with the appropriate strain obtained from a glycerol freeze, an overnight culture or a fresh single colony on an agar plate, using a flame sterilised nichrome wire loop. The plates were incubated at 37 ºC for 12 - 16 hours [14].

8.2.8 STRAIN STORAGE

Glycerol stocks were used to store strains. These were prepared by centrifuging 1.5 mL of overnight culture at 2655 *g* for 5 minutes in a screw top cryo-tube. The supernatant was aspirated and 700 µL of LB plus 300 µL of 50 % glycerol was added [14]. The pellet was resuspended gently and the tube stored at -80 ºC.

8.2.9 COMPETENT CELL PREPARATION

Competent cells were prepared using the calcium chloride technique [14]. A starter culture was prepared through the inoculation of 3 mL of LB broth containing the appropriate antibiotics. This was incubated at 37 ºC for 12 hours with shaking (180 rpm). A 1/50 volume of this culture was used to inoculate 50 – 200 mL of fresh media. This was incubated at 37 ºC with shaking at 180 rpm until the A_{600} was approximately 0.8. The culture was incubated on ice for 15 minutes then harvested through centrifugation at 2039 *g* at 4ºC for 10 minutes. The cells were resuspended in a 1/4 volume of ice cold 100 mM $CaCl_2$ then centrifuged as before. A second wash was performed in a 1/8 volume of ice cold 100 mM $CaCl_2$ then the culture was centrifuged again. The cells were finally resuspended in a 1/15 volume of ice cold 100 mM $CaCl_2$ and 50 µL aliquots were either used immediately in transformation, or stored at -80 ºC.

8.2.10 TRANSFORMATION

An aliquot of between 1 µL and 5 µL containing approximately 5 ng of DNA was added to 50 µL competent cells and incubated on ice for 30 minutes. The cells were heat shocked for 45 seconds at

42ºC and then incubated on ice for 2 minutes. 1 mL of SOC media was added and the cells elaborated for an hour at 37 ºC with gentle shaking. The cells were harvested by centrifugation at 2655 *g* for 1 minute. All but ~100 µL of the supernatant was aspirated. The cells were resuspended in the remaining supernatant and spread on LB plates containing the appropriate selection media and supplements [14]. The plates were incubated overnight at 37 ºC. Two controls were carried out concurrently with each experimental transformation. One contained no plasmid DNA. The other was carried out using the control plasmid pUC 18 or pWhitescript.

8.2.11 PLASMID PREPARATION

Solution 1:	50 mM Glucose
	10 mM EDTA
	25 mM Tris.HCl pH8.0
Solution 2:	0.2 M NaOH
	1 % SDS
Solution 3:	3 M Potassium acetate pH 4.8
	5 M Glacial acetic acid

A single bacterial colony was used to inoculate 3 mL of LB broth containing the appropriate antibiotics and culture was incubated at 37 ºC with shaking at 180 rpm overnight. 1.5 mL of culture was transferred to a centrifuge tube and spun at 10621 x *g* at 4 ºC for 5 minutes. The supernatant was removed by aspiration and the pellet was resuspended in 300 µL of ice-cold solution 1 then incubated on ice for 2 minutes. 300 µL of ice-cold, freshly prepared solution 2 was added and the tube was mixed by gentle inversion. The tube was incubated on ice for 2 minutes and then 300 µL of solution 3 was added and mixed by gentle inversion to produce a white precipitate. The tube was centrifuged at 15294 x *g* for 5 minutes and 850 µL of the supernatant was retained and mixed with 650 µL isopropanol. This was inverted to mix and then centrifuged at 15294 x *g* for 10 minutes to yield as small white pellet. The supernatant was aspirated and the pellet re-spun in the presence of 70 % ethanol. The ethanol was removed and the pellet left to dry at 37 ºC then resuspended in 80 µL sterile distilled water [14]. If further purification was required, a Purelink Quick Plasmid Miniprep kit (Invitrogen) was used.

8.2.12 RESTRICTION DIGESTION

The plasmid pJG001 and those arising from the point mutation of pJG001 were digested with *Eco*R I, *Hin*d III, or *Bam* HI (cutting at positions 701, 719 and 689 respectively) in a total of 10 μL [11]. The pET M11 plasmids were digested with *Eco*R V and *Xho* I (cutting at positions 2233 and 158 respectively). pET 151/D-TOPO plasmids were digested with EcoR V (cutting at positions 545 and 4782) [7, 8]. The reaction mixtures contained 5 μL plasmid DNA (~ 200 ng DNA), 1 μL restriction enzyme, 1 μL 10 X reaction buffer, and 3 μL dH_2O [14].

8.2.13 AGAROSE GEL ELECTROPHORESIS

Loading dye (10 x):	30 % v/v Glycerol
	0.25 % w/v Bromophenol blue
	0.25 % w/v Xylene cyanol
TAE buffer:	40 mM Tris-acetate pH 8.0
	1 mM EDTA

Plasmid DNA was run on a 1 % (w/v) agarose gel in conjunction with Hyperladder I which provided standards ranging between 200 and 10 000 bp. The gel was made up in a volume of 30 mL TAE buffer and heated until all the agarose was dissolved. Once at ~50 ºC, the gel was poured into a casting tray and allowed to set before the comb was removed and it was placed in a gel tank containing TAE buffer. The plasmid DNA was mixed with the loading buffer and loaded onto the gel alongside the Hyperladder I. Electrophoresis was conducted at 90 V for 1 to 1.5 hours or until the bromophenol blue band neared the bottom of the gel. The gel was stained with ethidium bromide for 20 minutes then visualised under a 302 nm UV light source [14].

8.2.14 PCR SITE-DIRECTED MUTAGENESIS

Site-directed mutagenesis was carried out as per the instructions in the Stratagene QuikChange site-directed mutagenesis manual [3].

Primer design

The Stratagene QuikChange mutagenesis kit makes use of synthetic mutagenic primers designed to introduce the desired mutation into the plasmid during PCR. The primers were designed as per the guidelines in the Stratagene QuikChange mutagenesis kit manual. Both primers were designed to contain the desired mutation and to anneal to the same sequence on opposite strands of the plasmid. The primers were between the prescribed 25 and 45 bp in length, with a melting temperature (T_m) of ≤ 78 ºC according to equation provided in the kit instruction manual. The desired mutations were in the middle of the primers with 10 – 15 bases of correct sequence either side. The primers were designed with a GC content of at least 40 % and with one or more C or G bases at the termini (figure 8.1). The primers were purified by Invitrogen using PAGE.

Parent strand	5’	GCC ATT AGC CTG ACG CAG CGC TTC AAT GAC AG	3’
Primer 1	5’	GCC ATT AGC CTG ACG CTG CGC TTC AAT GAC AG	3’
Primer 2	5’	CGG TAA TCG GAC TGC GAC GCG AAG TTA CTG TC	3’

Q90L

Figure 8.1 – The Q90L primer showing the position of the sequence alteration at position 90 (red). These primers are 32 bases long, has a T_M of 78 ºC in 1 M Na^+ and 57ºC in 50 mM Na^+ and a GC content of 56 %. Thermal denaturation temperatures were calculated using equation 8.1.

Equation 8.1 $T_M = 81.5 + 0.41(\%\ GC) - 675/N - \%\ mismatch$

The primers introducing the other mutations to the *E. coli* DHDPS sequence were designed in an analogous way.

Thermal cycling reaction conditions

PCR reactions were prepared using 5 and 50 ng of template DNA. A control reaction was performed using pWhitescript with mutagenic primers provided. The pWhitescript plasmid contains a point mutation that generates a stop codon in the lacZ gene. The primers provided replaced the stop codon, restoring the activity of the galactosidase enzyme coded by lacZ [3]. This allows the mutational efficiency of the reaction to be gauged through the assessment of the number of blue versus white colonies present on IPTG/X-Gal plates [14]. Mutagenesis reactions were prepared

using plasmids pJG001 or pSD001 as the template DNA. The plasmids were prepared as previously described (section 8.2.11). The DNA concentrations were determined spectrophotometerically by reading the absorbance at 280 nm. The reactions were performed in thin walled PCR tubes using a thermocycler equipped with a "hot top". The cycling parameters are summarised in table 8.3.

Step	Temperature (ºC)	Time (minutes)
1	95	2
2	95	0.5
3	55	1
4	68	13
5	Steps 2 – 4 repeated 15 times	

Fig 8.3 – The cycling parameters for the PCR reaction used in site directed mutagenesis [3].

Agarose gel electrophoresis was used to confirm amplification of the template plasmid in the presence of the mutagenic primers. The template plasmid and Hyperladder I were run in tandem.

Template digestion

The template DNA was digested through the addition of 1 μL (10 units) of *Dpn* I to each of the amplification reactions. *Dpn* I is a enzyme specific for methylated and hemimethylated DNA. This digests the parental DNA leaving the mutated cDNA intact. The enzyme was thoroughly mixed with the reaction mixture and incubated at 37 ºC for 1 hour.

Transformation of competent cells

E. coli XL1 Blue or BL21 (DE3) competent cells were thawed on ice and 1 μL of mutated DNA was added to the 50 μL aliquots. Following this, the protocol described in section 8.2.10 was followed.

DNA sequencing

DNA sequencing was carried out at Canterbury University using an ABI 3100 sequencer (Applied Biosystems). The plasmids to be sequenced were prepared to the highest quality, quantified spectrophotometrically, and sequenced using BigDye, buffer 6 (Applied Biosystems) and T7 and T7R primers for sequencing pJG001 or T7 and TopoR primers (Geneworks) for sequencing the pET plasmids.

8.3 GENERAL BIOCHEMICAL METHODS

All proteins were maintained at 4 ºC or on ice during experiments and stored in 20 mM Tris.HCl pH 8.0 (at 4 ºC) at -20 ºC. pH measurements were carried out using an UltraBasic UB -10 pH/mV meter (Denver Instruments). Centrifugation was carried out in an Eppendorf centrifuge 5810 R at 4ºC. For small volumes (1.5 mL or weighing ≤ 3.75 g) at up to 20817 *g* (14000 rpm) the Eppendorf F-45-30-11 rotor was used. For moderate volumes (50 mL or weighing ≤ 125 g) at up to 18353 *g* (12000 rpm) the Eppendorf F-34-6-38 rotor was used. For large volumes (500 mL or weighing ≤ 1.4 kg) at up to 3220 *g* (4000 rpm) the Eppendorf A-4-81 rotor was used. Chromatography columns were run using a UPC-900 ÄKTA FPLC (GE Healthcare) which was loaded using a Gilson Minipuls M312 peristaltic pump. In addition to this the BioRad Profinia was sometimes used. The fractions were collected with a Frac-950 fraction collector (GE Healthcare). PAGE gels were run using an Invitrogen ZOOM Dual Power unit using NuPage 4 – 12 % Bis-Tris gels (Invitrogen).

8.3.1 DETERMINATION OF PROTEIN CONCENTRATION

The standard Bradford assay was used to determine protein concentration [16]. For a standard spectrophotometer assay 800 µL of diluted protein was added to 200 µL of Bradford reagent (BioRad) and mixed thoroughly as per the instructions. The solution was incubated at room temperature (18 – 20 ºC) for 10 minutes and then the absorbance was measured at 595 nm against a blank consisting of 800 µL of sample buffer and 200 µL Bradford reagent. Protein concentrations were determined from a standard curve constructed using bovine serum albumin (BSA). All measurements were carried out in duplicate or triplicate. For assays carried out in the FLUOstar Optima plate reader (BMG Labtech) the Bradford reagent was diluted 1/4 with sterile distilled water. 10 µL of diluted protein sample was pipetted into a Microtest 96 well plate (Sarstedt) and 200 µL of diluted Bradford reagent added. The samples were mixed well and incubated at room temperature (18 – 20 ºC) for 10 minutes. The absorbance at 595 nm was measured and the protein concentrations of the samples determined from a standard curve constructed using 0 – 10 µg/mL BSA. All measurements were carried out in duplicate or triplicate.

Alternatively a Nanodrop ND 1000 Spectrophotometer (Thermo Scientific) was used to measure absorbance at 280 nm [17] Using the extinction coefficient (ε) as predicted by "Protein Calculator" (http://www.scripps.edu/~cdputnam/protcalc.html), the protein concentration was calculated.

8.3.2 SODIUM DODECYL SULFATE POLYACRYLAMIDE GEL ELECTROPHORESIS (SDS PAGE)

Electrophoresis tank buffer:	MOPS Buffer (containing 20 mM MOPS, 2 mM sodium acetate and 1 mM EDTA, pH 8.0) [14]
Coomassie blue stain:	1 % w/v Coomassie blue 50 % v/v Methanol 10 % v/v Glacial acetic acid
De-stain:	5 % v/v Methanol 10 % v/v Glacial acetic acid

For analysis of protein samples, pre-prepared NuPage 4 – 12 % Bis-Tris gels (Invitrogen) were placed into Xcel SureLock Mini-Cells and surrounded by >500 mL MOPS buffer containing 500 µL NuPage antioxidant. Between 10 and 20 µg of total protein was mixed with 2.5 µL NuPage LDS sample buffer and 1 µL NuPage reducing agent and the volume made up to 10 µL. Samples were heated at ~100 ºC for 5 minutes, centrifuged briefly and then loaded on the gel. Electrophoresis was conducted at 200 V for 40 - 50 minutes or until the bromophenol blue band neared the bottom of the gel. The gel was subsequently removed from the plastic casing, stained with Coomassie blue stain for 1 hour with shaking at ~55 rpm, and de-stained overnight. The gels were imaged using a Chemi Genius2 Bio Imaging System (Syngene).

8.3.3 PREPARATION OF DIALYSIS TUBING

Dialysis tubing was boiled for 10 minutes in a solution of 2 % (w/v) sodium bicarbonate and 0.05 % (w/v) EDTA. Following this, the tubing was boiled for an additional 10 minutes in two changes of deionised water. The tubing, once cool, was stored in a solution containing 0.02 % (w/v) sodium azide.

8.4 OVER-EXPRESSION AND PURIFICATION OF WILD-TYPE AND MUTANT DHDPS [18], DHDPR [6] AND TEV PROTEASE [13].

Buffer A: 20 M Tris.HCl (pH 8.0 at 20 ºC)

Buffer B: 50 mM Na_2PO_4 (pH 8.0), 300 mM NaCl, 30 mM imidazole

8.4.1 GROWTH OF *E. COLI* XL1 BLUE AND *E. COLI* BL21 (DE3)

Cultures of *E. coli* XL1 Blue or *E. coli* BL21 (DE3) containing the appropriate plasmids (coding for DHDPS) and were streaked onto LB agar plates supplemented with the appropriate antibiotics and incubated for 12 – 16 hours at 37 ºC. The *E. coli* XL1 Blue strain was used to express the non-polyhistidine tagged variants of DHDPS. The *E. coli* BL21 (DE3) strain was used to express those proteins that contained a polyhistidine motif. A single colony was used to inoculate a 3 - 5 mL LB broth, which was incubated overnight at 37 ºC with shaking (180 rpm). 1 mL aliquots from this overnight culture were used to inoculate flasks containing 1 L of LB or auto-induction broth (ZYM5052) supplemented with the appropriate antibiotics. These flasks were incubated at 37 ºC for 15 hours when inoculated with *E. coli* XL1 Blue. The cultures containing the *E. coli* BL21 (DE3) were induced with 1 mM IPTG after 8 hours and harvested 12 hours from inoculation. Those cultures grown in ZYM5052 were incubated at 20 ºC or 37 ºC for 12 hours prior to harvesting. The cells were harvested by centrifugation at 3220 *g* for 10 minutes and the supernatant discarded. The XL1 Blue cells were washed by re-suspending the pellet in 0.25 volumes of buffer A by gentle pipetting. The BL21 (DE3) cells were washed using buffer B. The cells were re-harvested through centrifugation at 3220 *g* for 10 minutes and the supernatant discarded. The pellet was re-suspended in a small volume of buffer A or buffer B as appropriate and centrifuged as before. The final XL1 Blue or BL21 (DE3) cell pellet was re-suspended in 30 mL of buffer A or buffer B respectively. Typically 1 L of inoculated LB broth yielded between 2 – 4 g of XL1 Blue cells (wet-weight) or 3 – 5 g of BL21 (DE3) cells. ZYM5052 media typically yielded 3 – 5 g cells (wet weight).

8.4.2 GROWTH OF *E. COLI* BL21 (DE3) pET M11 dapB

Cultures *E. coli* BL21 (DE3) containing pET M11 dapB plasmid (coding for DHDPR [6]) were prepared as described using the techniques described in section 8.4.1. The final cell pellet was re-suspended in 30 mL of buffer B. Typically 1 L of inoculated media yielded 3 – 5 g cells.

8.4.3 GROWTH OF *E. COLI* BL21 (DE3) pRK793

Cultures *E. coli* BL21 (DE3) containing the pRK 793 plasmid (coding for TEV protease [13]) were prepared as described using the techniques described in section 8.4.1. The final cell pellet was re-suspended in 30 mL of buffer B. Typically 1 L of inoculated media yielded 3 – 5 g cells.

8.4.4 PREPARATION OF CRUDE CELL FREE EXTRACTS

The crude cell free extract was prepared by sonication using a Sonics vibracell sonicator. The harvested cells were exposed to ultra-sonication of 3 second bursts followed by 10 second pauses. This was carried out until the cells had received approximately 5000 J [11]. Following sonication the cell extract was centrifuged at 12745 *g* for 30 minutes. The supernatant was retained and stored on ice while the pellet was resuspended in 15 mL buffer A and subjected to sonication as described above. The supernatants were pooled.

8.4.5 ION EXCHANGE CHROMATOGRAPHY (IEC) [19, 20]

Buffer C: 20 mM Tris.HCl, (pH 8.0 at 20 ºC), 1 M NaCl

The pooled supernatants from the sonication step were loaded into the Superloop of the ÄKTA FPLC using a Gilson M312 peristaltic pump. The Q Sepharose ion exchange column (36.5 mL) was

equilibrated with 5 bed volumes of buffer A. The sample was loaded onto the column and the enzyme was eluted with a 0 – 1 M NaCl gradient in buffer A (using buffer C) over 4 column volumes. DHDPS eluted at approximately 30 mS/cm. The column was washed with 3 bed volumes of buffer C [11, 19]. The eluted fractions (1 – 2 mL) were tested for activity using the *o*-aminobenzaldehyde assay as described in section 8.4.8. Those exhibiting activity were pooled and used immediately or stored at – 20 ºC.

8.4.6 HYDROPHOBIC INTERACTION CHROMATOGRAPHY (HIC) [19, 21]

Buffer D: 20 mM Tris.HCl (pH 8.0 at 20 ºC), 0.5 M $(NH_4)_2SO_4$

Ammonium sulfate was added to the pooled, dialysed fractions from the ion exchange step to a final concentration of 0.5 M. This was loaded into the Superloop of the ÄKTA FPLC. The phenyl-Sepharose column (23.5 mL) was equilibrated with 3 bed volumes of buffer D. The sample was loaded onto the column and the enzyme was eluted with a 0.5 – 0 M ammonium sulfate gradient over 4 column volumes or stepwise over 0 column volumes. The column was washed with 3 bed volumes of buffer A [11, 19]. The eluted fractions were tested for activity using the *o*-aminobenzaldehyde assay as described in section 8.4.8. The active fractions eluted between 0.2 and 0.1 M ammonium sulfate. These were pooled then dialysed against 100 volumes of buffer A overnight at 4 ºC or run through the a HiPrep 26/10 desalting column using buffer A.

8.4.7 AFFINITY CHROMATOGRAPHY [22-24]

Buffer E: 50 mM Na_2PO_4 (pH 8.0), 300 mM NaCl, 300 mM imidazole

The crude sample was loaded onto a 5 mL His-Trap FF crude column that had been pre-equilibrated with 3 bed volumes of buffer B. The sample was loaded into the Superloop of the ÄKTA FPLC then loaded onto the column. The enzyme was eluted using buffer E and the fractions exhibiting UV absorbance were pooled and dialysed against buffer A. The eluted fractions of DHDPS were tested for activity using the *o*-aminobenzaldehyde assay as described in section 8.4.8. Those

exhibiting activity were pooled and used immediately or stored at –20 ºC. The eluted fractions of DHDPR and TEV protease exhibiting absorbance at 280 nm were pooled.

8.4.8 *o*-AMINOBENZALDEHYDE ASSAY [25]

The assay was carried out in a 96 well plate in which 150 µL of the assay mixture (containing 180 mM HEPES, 20.65 mM *o*-aminobenzaldehyde, 18 mM pyruvate, and ~1.8 mM (*S*)-ASA) was added to 1-5 µL of each protein fraction. This was incubated at 37 ºC for ~10 minutes and then the reaction was quenched through the addition of 100 µL TCA (10 % w/v). The active fractions exhibited a purple colouration and the corresponding fractions were pooled. The negative control reaction contained no enzyme. The positive control reaction contained enzyme from an earlier purification, the activity of which had been verified.

8.4.9 TEV PROTEASE CLEAVAGE OF POLYHISTIDINE TAGS

The cleavage of the polyhistidine tag from the DHDPS variants was carried out using TEV protease [13] (prepared as per section 8.4.3 and 8.4.7). The TEV protease and pET 151/D-TOPO polyhisitide tagged *E. coli* DHDPS variants were incubated together (with the TEV protease at a concentration of 1 % w/w of the DHDPS variant) in the presence of 50 mM Tris.HCl (pH 8.0), 0.5 mM EDTA and 1 % v/v 0.1 M DTT [13]. The reaction mixture was incubated at 4 ºC for 48 hours. The cleaved DHDPS variants were harvested by applying the reaction mixture to a 5 mL His-Trap FF crude column and collecting the flow-through and 2 column washes performed with buffer B. The uncleaved DHDPS and the TEV protease were eluted from the column using buffer E. The resulting protein was dialysed against buffer A or run through a HiPrep 26/10 desalting column using buffer A.

8.5 KINETIC ANALYSIS OF WILD-TYPE DHDPS AND ITS VARIANTS

Two-substrate steady-state rate analysis was employed [26]. DHDPS activity was measured through a coupled assay to DHDPR (prepared as per section 8.4.2), as previously described by Coulter [27].

8.5.1 INITIAL RATE DETERMINATION AND KINETIC ANALYSIS

The kinetic assays, following the techniques of Dobson *et al.* [28, 29], were performed in 100 mM HEPES buffer (pH 8.0) at 30 ºC. The temperature was controlled using a circulating water bath. Buffer was used as a blank. The concentration of DHDPR was kept in excess of that of DHDPS ensuring that the reaction catalysed by DHDPS was the rate limiting step. Following the addition of DHDPR and components other than DHDPS and (*S*)-ASA, the samples were allowed to equilibrate to 30 ºC for 10 minutes. The DHDPS was then added to the reaction mixture, immediately followed by the (*S*)-ASA. The reaction was monitored by measuring the change in wavelength at 340 nm on a Hewlett Packard 8452A diode array spectrophotometer with a circulating water bath to maintain constant temperature. This change in absorbance can be correlated to the consumption of NADPH by DHDPR. The concentration of DHDPS was kept to a level such that the initial rate observed was between 0.5 and 10 mAU/sec.

Kinetic parameter determination

The K_m apparents and k_{cat} were determined through the variation of the substrate concentrations. Substrate concentrations were varied between 0.1 and 10 times the K_m apparents. The velocities were measure in at least triplicate and the values obtained were typically within 10 % error.

	Initial concentration	Volume (µL)	Final concentration
HEPES (pH 8.0, 25 °C)	100 (mM)	780	78 (mM)
NADPH	3.24 (mM)	50	0.162 (mM)
DHDPR	~ 1 mg/mL	50	~ 0.005 mg/mL
DHDPS	0.3 – 2.5 mg/mL	20	0.0006 – 0.005 mg/mL
(*S*)-ASA	1 – 160 (mM)	50	0.05 – 8 (mM)
Pyruvate	1 – 160 (mM)	50	0.05 – 8 (mM)
Total		1000	

Table 8.4 – The experimental design for the coupled assay used for the analysis of the K_m apparents of DHDPS and its variants. The amount of DHDPR was kept such that it did not limit the initial rate of the reaction. Typically ~ 50 µg of DHDPS was added to each assay.

The K_m apparents were estimated by varying the concentration of one substrate whilst holding the other at its maximum concentration. The (*S*)-ASA concentration was kept constant at 2 mM (final

concentration) whilst the pyruvate concentration was varied between 0.05 and 2 mM (final concentration). Likewise, the pyruvate was kept at 2 mM whilst the (*S*)-ASA concentration was varied between 0.5 and 2 mM (final concentrations). The K_m apparents were estimated from a data set of a minimum of 7 substrate concentrations, each of which was measured in at least triplicate. Initial rate data were fitted using Origin 7 (OriginLab, Northhampton, MA). The data were fitted to a hyperbolic function and the K_m apparents and the k_{cat} parameters calculated.

8.6 BIOPHYSICAL METHODS

8.6.1 ANALYTICAL GEL PERMEATION LIQUID CHROMATOGRAPHY [30]

Buffer F: 20 mM Tris.HCl (pH 7.0 at 20ºC), 100 mM NaCl

Analytical gel permeation liquid chromatography was performed using Sephadex G-200 (10/300 GL prepacked) superfine column with a 24 mL bed volume. Prior to running the protein samples the column was equilibrated with buffer F. Calibration of the column was carried out using yeast alcohol dehydrogenase, BSA, ovalbumin and ribonuclease A (RNase A) [11]. 300 µL of the protein (at 0.5 mg/mL) was loaded onto the equilibrated column and eluted with buffer F. The absorbance of the elutant was monitored at 280 nm and these data were used to calculate elution volumes (V_e) by zonal analysis [30]. The resulting data were used to either construct a calibration curve of the V_e versus the $\log_{10}$ of the known molecular weight or to ascertain the molecular weight of the unknown samples.

8.6.2 ANALYTICAL ULTRA-CENTRIFUGATION (AUC)

Analytical ultra-centrifugation experiments [31, 32] were carried out on a Beckman RFT XL-I analytical ultra-centrifuge equipped with UV/Vis scanning optics and an An-60 Ti 4-hole rotor or 8-hole rotor. The protein samples were desalted into buffer F (using a HiPrep 26/10 desalting column) and loaded into 12 mm double sector cells with quartz windows.

Sedimentation velocity analytical ultra-centrifugation

Samples with an initial protein concentration of 0.1 mg/mL (380 µL) and a reference containing buffer F only (400 µL) were centrifuged at 40 000 rpm at 18 ºC and data were collected at 235 nm

every 10 minutes without averaging and with a step size of 0.03 mm. Sedimentation velocity data were fitted to a continuous size-distribution model using the program SEDFIT [33]. The partial specific volume ($\bar{v}$) of the sample and the buffer density and viscosity of the reference solution were calculated using the program SEDNTERP [34].

8.6.3 DIFFERENTIAL SCANNING FLUORIMETRY

The thermal denaturation temperatures of the proteins of interest were assessed using differential scanning fluorimetry (DSF), carried out in a BioRad IQ5 ICycler Real Time PCR machine [35]. The stability of each protein at 0.5 mg/mL was assessed in 100 mM phosphate buffer at pH 2, pH 4.5, pH 7, pH 8.5 and pH 11 with 0 mM, 40 mM, 100 mM and 500 mM NaCl resulting in a total of 20 conditions. 25 µL aliquots of 50 x SYPRO Orange (Molecular Probes) were added to 75 µL of test solution in each well of a thin walled 96 well PCR plate (BioRad) [36, 37]. Each assay was performed three times in the same run. The control samples contained water instead of protein. The plates were sealed with Microseal B adhesive sealer (BioRad) and heated in the ICycler from 20 ºC to 94.1 ºC in 20 second increments of 0.3 ºC. Changes in the fluorescence of the wells were measured using the attached camera with an excitation filter at 490 nm and an emission filter at 575 nm. Analysis of the resultant data was performed. The point at which the protein thermally denatured was assessed by ascertaining the peak of the derivatised data. [35, 36, 38-41]

Statistical analysis of the DSF data was carried out using "R" (http://www.r-project.org/) [42]. The underlying error distribution was non-normal so a generalised linear model with a Poisson error and a log link function was used. Melting temperature was the response, and protein, pH level and salt concentration were the categorical factors. There was no evidence of over-dispersion, and significance tests were made using a Chi-square statistic.

8.6.4 CIRCULAR DICHROISM SPECTROSCOPY

CD spectra were recorded at the University of Canterbury on a Jasco J-815 CD spectrophotometer at 20 ºC in 20 mM phosphate buffer (pH 7.0) containing 40 mM NaCl in a 0.1 mm quartz cell. The spectrophotometer was blanked with the buffer (20 mM phosphate buffer (pH 7), 40 mM NaCl).

The slit bandwidth used was 1 nm with a step size of 0.5 nm. Measurements were taken with a signal averaging time of 1.0 second.

CD melts were performed under the same conditions between 20 ºC and 80 ºC with a temperature slope of 1 ºC/min. The wavelength was monitored at 220 nm with a data pitch of 0.5 ºC and a delay time of 30 seconds. Wavelength scans (185 – 250 nm) were performed at every 10 ºC interval within the range of 20 ºC to 80 ºC. Analysis of secondary structural composition was carried out using CDPro (http://lamar.colostate.edu/~sreeram/CDPro/) [43].

8.6.5 AMORPHOUS AGGREGATION

Temperature dependent aggregation was measured by monitoring the absorbance at 340 nm [44]. The assays were performed at pH 2, pH 4.5, pH 7, pH 8.5 and pH 11 100 mM phosphate buffer supplemented with 0 mM, 20 mM, 40 mM, 100 mM and 500 mM NaCl resulting in a total of 25 conditions. The 200 µL assays were performed at 0.5 mg/mL protein and were carried out at 60 ºC in a 96 well microplate (Sarstedt) sealed with sealing tape (NUNC). A FLUOstar Optima platereader (BMG Labtech) was used to assess aggregation by monitoring the absorbance at 340 nm every 120 seconds for 50 cycles. The well scanning was performed in an orbital manner with a scanning diameter of 3 mm. Five flashes were performed per cycle and the microplate was shaken for 3 seconds preceding each reading. The controls performed were carried out in the absence of the protein and the resultant data were used as blanks. Each assay was performed 3 times with 2 replicates performed each time. The increase in absorbance was fitted to a sigmoid (logistic) function (equation 8.1) using Origin (OriginLab, Northhampton,MA) from which the maxima and the rate of amorphous aggregation ($Agg_{1/2}$) were calculated (section 2.9) [39, 45].

Equation 8.1 $$y = \frac{A_1 - A_2}{1 + (x/x_o)^p} + A_2$$

The initial value is represented by A_1, the final value by A_2. The midpoint of the curve can be calculated and dx is a time constant.

8.6.6 β-SHEET-SPECIFIC AGGREGATION

The β-sheet-specific aggregation was assessed at pH 2, pH 4.5, pH 7, pH 8.5 and pH 11 100 mM phosphate buffer supplemented with 0 mM, 40 mM, 100 mM and 500 mM NaCl resulting in a total of 20 conditions. The 200 µL assays were performed at 0.5 mg/mL (0.016 mM) protein in the presence of 0.02 mM thioflavin T (ThT) and were carried out at 60 ºC in a 96 well microplate (BMG LabTech) sealed with sealing tape (NUNC). A FLUOstar optima platereader (BMG Labtech) was used to assess the ThT induced fluorescence [46-49]. The excitation and emission filter were at wavelengths of 450 and 485 nm respectively. 250 cycles were performed at a cycle time of 1027 seconds. The well scanning was performed according to the parameters described in section 8.6.5. The controls performed were carried out in the absence of the protein and the resultant data were used as blanks. Each assay was performed 3 times with 2 replicates performed each time. The resulting increase in fluorescence was fitted to a sigmoid (Boltzmann) function (equation 8.1) using Origin (OriginLab, Northhampton,MA) from which the maxima and the rate of β-sheet specific aggregation ($\beta\text{-}Agg_{½}$) were calculated (section 2.9) [45]. The incubation of insulin at pH 1.7 in the presence of ThT was carried out as a positive control alongside the DHDPS assays. These wells contained 0.1 mM insulin in the presence of 0.02 mM ThT. These wells consistently demonstrated fluorescence readings in excess of 65 000 RFU; the detection limit of the FLUOstar optima platereader. The effect of salt on the fluorescence of ThT in the presence of insulin was carried about as above but in the presence of 0 mM, 10 mM, 40 mM, 100 mM and 500 mM NaCl. The assays were either carried out over a continuous time course, adding the ThT at time zero, or by adding ThT to samples at various time periods.

8.7 X-RAY CRYSTALLOGRAPHY

Mother liquor 2 M K_2HPO_4 pH 10.0

Crystallisation experiments were performed using the hanging-drop vapour-diffusion method as utilised for DHDPS in prior studies [19, 29, 50]. 1 mL aliquots of the mother liquor were placed in greased VDX 24 well plates (Hampton Research). The hanging drops were prepared by placing 2.5, 2.75 and 3.0 µL of filtered protein (~10 mg/mL) onto 22 mm siliconised cover slides (Hampton Research). To these drops, 0.9, 1, 1.1, 1.2, or 1.3 µL of the mother liquor and 0.6 µL NOG (6 %

w/v) was added. The total volume of the drop was made up to 4.9 µL through the addition of sterile distilled water. The cover slides were quickly inverted and placed over wells containing the mother liquor and sealed through the application of gentle pressure. Crystals were formed over 3 – 4 days. Prior to mounting and freezing in liquid nitrogen, the crystals were immersed in 20 % v/v glycerol, 2 M K_2HPO_4 pH 10.0.

X-ray data collection

Intensity data were collected at (110 K using a RAxisIV++ image-plate detector coupled to a Rigaku Micromax 007 X-ray generator operating at 40 kV and 20 mA. Diffraction data sets were processed and scaled using the program CrystalClear [50].

Structural determination and refinement

Structures were solved using AMoRe [52, 53], using the *E. coli* DHDPS monomer (PDB accession code 1yxc) [29] as the search mdel. Refinement was achieved using REFMAC5 [52, 54] with manual mode corrections using the program WinCoot (http://www.chem .gla.ac.uk/~bernhard/coot/ wincoot-download.html) [55]. Placement of the solvent molecules was made by the program ARP [51, 52]. Water molecules with a B-factor of ≥ 50 were removed. Structure quality was assessed using Procheck [52, 56].

Structural figures were produced using PyMOL [57]. Figures showing wild-type *E. coli* DHDPS structures were produced using the coordinates (PDB accession code 1yxc) of the structure solved by Dobson *et al.* (2005) [29].

8.8 CONFIRMATION OF AMYLOID

8.8.1 TRANSMISSION ELECTRON MICROSCOPY (TEM)

The protein samples were incubated at 60 ºC for three hours in 100 mM phosphate buffer of a pH and salt concentration, which yielded the highest ThT fluorescence. Prior to application to the TEM grids, the samples were diluted 1:20 to 1:100 with nanopure water.

Initially the formvar coated TEM grids (Electron Microscopy Sciences) were washed with 3 µL nanopure water. Following this 3 µL of protein sample was applied and incubated at room

temperature for 90 seconds. Subsequently the grids were washed with 3 μL nanopure water and negatively stained with 1 % w/v uranyl acetate (Electron Microscopy Sciences) for 45 seconds (adapted from Meehan *et al.* 2004 [58]). The grids were dried with filter paper between each application of solution. The samples were viewed under a JEOL 1200EX or a Morgagni 268D TEM.

8.8.2 X-RAY FIBRE DIFFRACTION

The stalks for X-ray fibre diffraction were prepared using the stretch-frame method [59]. Glass capillary tubes were cut to lengths of ~ 1cm and the ends coated in melted beeswax. The capillaries were placed with the coated ends opposite one another at a distance of approximately 2 mm and secured with plasticine. The DHDPS samples were prepared by incubating the protein for 3 hours at 60 ºC in the buffer which resulted in the highest ThT fluorescence. An aliquot of this suspension was placed between the coated ends of the capillaries and allowed to dry. During the drying process the capillaries were incrementally moved apart, stretching the sample and encouraging alignment of the fibrils.

8.8.3 FIBRIL PURIFICATION

Proteins were incubated at 60 ºC for three hours in 100 mM phosphate buffer of a pH and salt concentration, which yielded the highest ThT fluorescence. The aggregated samples were then digested with trypsin for 90 minutes at 37 ºC in 100 mM phosphate buffer (pH 7), 40 mM NaCl. The trypsin stock was made up in 1 mM HCl to a final concentration of 20 mg/mL and added to the protein samples at a ratio of 1:200 times the original protein concentration. The digested samples were centrifuged in a Beckman Coulter Optima – L 90K Ultracentrifuge with an NVT 65 rotor at 65000 rpm for 90 minutes at 4 ºC. The cell pellet was resuspended in dH_2O and examined by TEM and ThT.

8.9 REFERENCES

1 Roberts, S. J., Morris, J. C., Dobson, R. C. J. and Gerrard, J. A. (2003) The preparation of (*S*)-aspartate semi-aldehyde appropriate for use in biochemical studies. Bioorganic and Medicinal Chemistry Letters **13**, 265 - 267

2 Coulter, C. V., Gerrard, J. A., Kraunsoe, J. A. E. and Pratt, A. J. (1996) (*S*)-Aspartate semi-aldehyde: synthetic and structural studies. Tetrahedron **52**, 7127 - 7136
3 Stratagene (2006) Quikchange site-directed mutagenesis kit: instruction manual. http://www.stratagene.com/manuals/200516.pdf
4 Gerrard, J. A. (1992) DPhil thesis. Studies on dihydrodipicolinate synthase, Oxford University, Oxford
5 Coulter, C. V. (1997) PhD thesis. Studies in lysine biosynthesis, University of Canterbury, Christchurch
6 Devenish, S. (2007) Cloning *dap*A and *dap*B into pET vectors. Personal communication, Christchurch
7 Dümmler, A., Lawrence, A.-M. and de Marco, A. (2005) Simplified screening for the detection of soluble fusion constructs expressed in *E. coli* using a modular set of vectors. Microbial Cell Factories **13**, 34
8 Invitrogen (2004) Champion pET directional TOPO expression kits version G - Instruction manual.
9 Pearce, F. G., Dobson, R. C., Weber, A., Lane, L. A., McCammon, M. G., Squire, M. A., Perugini, M. A., Jameson, G. B., Robinson, C. V. and Gerrard, J. A. (2008) Mutating the tight-dimer interface of dihydrodipicolinate synthase disrupts the enzyme quaternary structure: towards a monomeric enzyme. Biochemistry **In press**
10 Dobson, R. C. J., Valegard, K. and Gerrard, J. A. (2004) The crystal structure of three site-directed mutants of *Escherichia coli* dihydrodipicolinate synthase: Further evidence for a catalytic triad. Journal of Molecular Biology **338**, 329 - 339
11 Griffin, M. D. W. (2005) PhD Thesis. Why is DHDPS a tetramer?, University of Canterbury, Christchurch
12 Norrander, J., Kempe, T. and Messing, J. (1986) Construction of improved M13 vectors using oligonucleotide-directed mutagenesis. Gene **26**, 101 - 106
13 Kapust, R. B., Tözsér, J., Fox, J. D., Anderson, D. E., Cherry, S., Copeland, T. D. and Waugh, D. S. (2001) Tobacco etch virus protease: Mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. Protein Engineering **14**, 993 - 1000
14 Sambrook, J. and Russell, D. W. (2001) Molecular cloning: A laboratory manual. Cold Springs Harbor Laboratory Press., New York
15 Studier, F. W. (2005) Protein production by auto-induction in high density shaking cultures. Protein Expression and Purification **41**, 207 - 234
16 Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical Biochemistry **72**, 248 - 254
17 Stoscheck, C. M. (1990) Quantitation of protein. Methods in Enzymology **182**, 50 - 68
18 Dobson, R. C. J., Gerrard, J. A. and Pearce, F. G. (2004) Dihydrodipicolinate synthase is not inhibited by its substrate, (*S*)-aspartate β-semialdehyde. Biochemical Journal **377**, 757 - 762
19 Mirwaldt, C., Korndorfer, I. and Huber, R. (1995) The crystal structure of dihydrodipicolinate synthase from *Escherichia coli* at 2.5 Å resolution. Journal of Molecular Biology **246**, 227 - 239
20 Yamamoto, S., Nakanishi, K. and Matsuno, R. (1988) Ion-exchange chromatography of proteins. M. Dekker, New York
21 Hemstrom, P. and Irqum, K. (2006) Hydrophobic interaction chromatography. Journal of Separation Science **29**, 1784 - 1821
22 Wilcheck, M. and Miron, T. (1999) Thirty years of affinity chromatography. Reactive and Functional Polymers **41**, 263 - 268

23 Cuatrecasas, P., Wilcheck, M. and Anfinsen, C. B. (1968) Selective enzyme purification by affinity chromatography. Proceedings of the National Academy of Sciences USA **61**, 636 - 643

24 Hengen, P. N. (1995) Purification of his-tag fusion proteins from *Escherichia coli*. Trends in Biochemical Sciences **20**, 285 - 286

25 Yugari, Y. and Gilvarg, C. (1965) The condensation step in diaminopimelate synthesis. The Journal of Biological Chemistry **240**, 4710 - 4716

26 Cornish-Bowden, A. (1999) Fundamentals of enzyme kinetics. Portland Press Ltd., London

27 Coulter, C. V., Gerrard, J. A., Kraunsoe, J. A. E. and Pratt, A. J. (1999) *Escherichia coli* dihydrodipicolinate synthase and dihydrodipicolinate reductase: Kinetic and inhibition studies of two putative herbicide targets. Pesticide Science **55**, 887 - 895

28 Dobson, R. C. J., Griffin, M. D. W., Roberts, S. J. and Gerrard, J. A. (2004) Dihydrodipicolinate synthase (DHDPS) from *Escherichia coli* displays partial mixed inhibition with respect to its first substrate, pyruvate. Biochimie **86**, 311 - 315

29 Dobson, R. C. J., Griffin, M. D. W., Jameson, G. B. and Gerrard, J. A. (2005) The crystal structures of native and (*S*)-lysine-bound dihydrodipicolinate synthase from *Escherichia coli* with improved resolution show new features of biological significance. Acta Crystallographica Section D, Biological Crystallography **D61**, 1116 - 1124

30 Winzor, D. J. (2003) Analytical exclusion chromatography. Journal of Biochemical and Biophysical Methods **56**, 15 - 52

31 Lebowitz, J., Lewis, M. S. and Schuck, P. (2002) Modern analytical ultracentrifugation in protein science: a tutorial review. Protein Science **11**, 2067 - 2079

32 Laue, T. M. and Stafford, W. F. (1999) Modern applications of analytical ultracentrifugation. Annual Review of Biophysics and Biomolecular Structure **28**, 75 - 100

33 Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. Biophysical Journal **78**, 1606 - 1619

34 Laue, T. M., Shah, B. D., Ridgeway, T. M. and Pelletier, S. L. (1992) Analytical ultracentrifugation in biochemistry and polymer science. In Analytical ultracentrifugation in biochemistry and polymer science., pp. 90 - 125, The Royal Society of Chemistry, Cambridge

35 Epps, D. E., Sarver, R. W., Rogers, J. M., Herberg, J. T. and Tomich, P. K. (2001) The ligand affinity of proteins measured by isothermal denaturation kinetics. Analytical Biochemistry **292**, 40 - 50

36 Steinberg, T. H., Haugland, R. P. and Singer, V. L. (1996) Applications of SYPRO orange and SYPRO red protein gel stains. Analytical Biochemistry **239**, 238 - 245

37 Steinberg, T. H., Jones, L. J., Haugland, R. P. and Singer, V. L. (1996) SYPRO orange and SYPRO red gel stains: One step fluorescent staining of denaturing gels for detection of nanogram levels of protein. Analytical Biochemistry **239**, 233 - 237

38 Niesen, F. H., Berglund, H. and Vedadi, M. (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. Nature Protocols **2**, 2212 - 2221

39 Vedadi, M., Niesen, F. H., Allali-Hassani, A., Fedorov, O. Y., Finerty, P. J. J., Wasney, G. A., Yeung, R., Arrowsmith, C., Ball, L. J., Berglund, H., Hui, R., Marsden, B. D., Norlund, P., Sundstrom, M., Weigelt, J. and Edwards, A. M. (2006) Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. Proceedings of the National Academy of Sciences USA **103**, 15835 - 15840

40 Senisterra, G. A., Hong, B. S., Park, H.-W. and Vedadi, M. (2008) Application of high-throughput isothermal denaturation to assess protein stability and screen for ligands. Journal of Biomolecular Screening **13**, 337 - 342

41 Stryer, L. (1965) The interaction of a napthalene dye with apomyoglobulin and apohemoglobin. A fluorescent probe of non-polar binding sites. Journal of Molecular Biology **13**, 482 - 495
42 Venables, W. N., Smith, D. M. and Team, R. D. C. (2008) An introduction to R. In Notes on R: A programming environment for data analysis and graphics.
43 Sreerame, N. and Woody, R. W. (2000) Estimation of protein secondary structure from circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. Analytical Biochemistry **287**, 252 - 260
44 Poon, S. and Rivers, R. C. (2005) Protein aggregation monitoring on the FLUOstar OPTIMA Microplate reader. http://www.bmglabtech.com/db_assets /applications/ downloads/ applications/ AN135-protein-aggregation-fluorescence-FLUOstar.pdf
45 Sabate, R., Gallardo, M. and Estelrich, J. (2003) An autocatalytic reaction as a model for the kinetics of the aggregation of β-amyloid. Biopolymers (Peptide Science) **71**, 190 - 195
46 Saeed, S. M. and Fine, G. (1967) Thioflavin-T for amyloid detection. American Journal of Clinical Pathology **57**, 588 - 593
47 Khurana, R., Coleman, C., Ionescu-Zanetti, C., Carter, S. A., Khrishna, V., Grover, R. K., Roy, R. and Singh, S. (2005) Mechanism of thioflavin T binding to amyloid fibrils. Journal of Structural Biology **151**, 229 - 238
48 Groenning, M., Norrman, M., Flink, J. M., van der Weert, M., Butrinsky, J. T., Schluckebier, G. and Frokjaer, S. (2007) Binding of thioflavin T in insulin amyloid fibrils. Journal of Structural Biology **159**, 483 - 497
49 LeVine III, H. (1993) Thioflavine T interaction with synthetic Alzheimer's disease β-amyloid peptides: Detection of amyloid aggregation in solution. Protein Science **2**, 404 - 410
50 Pflugrath, J. W. (1999) The finer things in X-ray diffraction data collection. Acta Crystallographica Section D, Biological Crystallography **55**, 1718 - 1725
51 Lamzin, V. S. and Wilson, K. S. (1997) Automated refinement for protein crystallography. Methods in Enzymology **277**, 269 - 305
52 Bailey, S. (1994) The CCP4 suite: programs for protein crystallography. Acta Crystallographica Section D, Biological Crystallography **50**, 760 - 763
53 Navaza, J. and Saludjian, P. (1997) AMoRe: An automated molecular replacement program package. Methods Enzmol. **276**, 581 - 594
54 Murshudov, G. N., Vagin, A. A. and Dodson, E. J. (1997) Refinement of macromolecular structures by maximum-likelihood method. Acta Crystallographica Section D, Biological Crystallography **53**, 240 - 255
55 Emsley, P. and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. Acta Crystallographica Section D: Biological Crystallography **60**, 2126 - 2132
56 Laskowski, R. A., Macarthur, M. W., Moss, D. S. and Thornton, J. M. (1993) Procheck: A program to check the sterochemical quality of protein structures. Journal of Applied Crystallography **26**, 283 - 291
57 DeLano, W. L. (2002) The PyMOL molecular graphics system. DeLano Scientific, San Carlos
58 Meehan, S., Berry, Y., Luisi, B., Dobson, C. M., Carver, J. A. and MacPhee, C. E. (2004) Amyloid fibril formation by lens crystallin proteins and its implications for cataract formation. The Journal of Biological Chemistry **279**, 3413 - 3419
59 Serpell, L. C., Fraser, P. E. and Sunde, M. (1999) X-ray fiber diffraction of amyloid fibrils. In Aggregate and Precursor Protein Structure: Aggregate Morphology (Nielsen, E. H., Nybo, M. and Svehag, S.-E., eds.), pp. 526 - 537, Academic Press

APPENDIX 1

SUPPLEMENTARY INFORMATION FOR CHAPTER 2

A1.1 DIFFERENTIAL SCANNING FLUORIMETRY

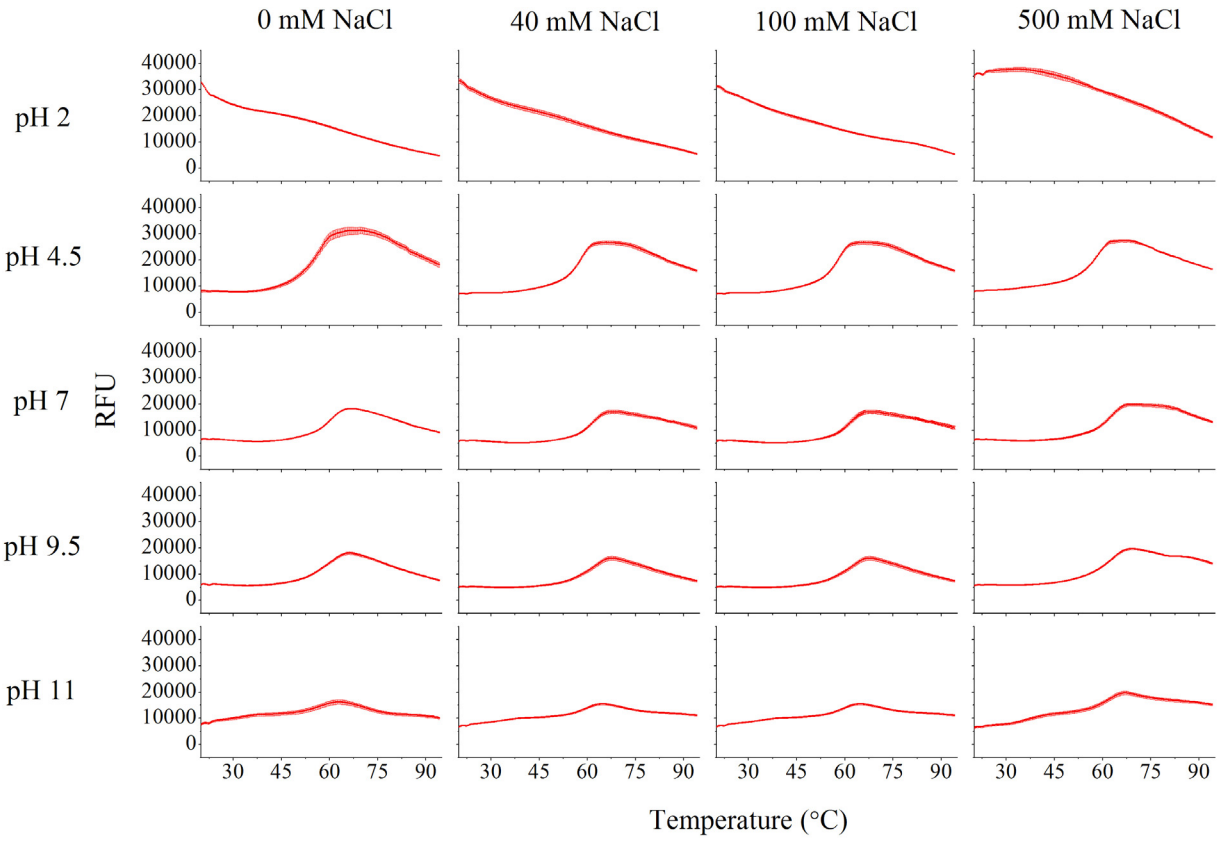


Figure A1.1 – The thermal denaturation of wild-type *E. coli* DHDPS in 100 mM phosphate buffer at pH 2, pH 4.5, pH 9.5 and pH 11 containing 0 mM NaCl, 40 mM NaCl, 100 mM NaCl and 500 mM NaCl as monitored by SYPRO orange fluorescence. The data plotted are the increase in fluorescence, as monitored by the BioRad IQ5.

pH	Salt concentration (mM)	Mean thermal denaturation temperature (°C)
4.5	0	57.8 ± 0.5
	40	57.7 ± 0.2
	100	57.9 ± 0.2
	500	57.7 ± 0.2
7	0	60.5 ± 0.3
	40	60.9 ± 0.3
	100	61.5 ± 0.1
	500	61.8 ± 0.2
9.5	0	60.8 ± 1.0
	40	61.9 ± 0.4
	100	61.9 ± 0.3
	500	62.2 ± 0.1
11	0	57.2 ± 0.3
	40	57.3 ± 0.4
	100	59.0 ± 0.0
	500	59.2 ± 0.1

Table A1.1 – The thermal denaturation temperatures (T_m) of wild-type *E. coli* DHDPS in 100 mM phosphate buffer at various pH and salt concentrations as monitored by SYPRO orange fluorescence and as assessed from the derivation of the fluorescence data. The errors were calculated by identifying the maximal peak of the derivatives of each of the three replicates. The temperature at which each maximum occurred was established and the mean and standard error of the mean calculated.

	Degrees of freedom	Deviance	Residual degrees of freedom	Residual deviance	P (> Chi)
Null			1055	2838.06	
Protein	21	1813.09	1034	1024.97	0
pH	3	338.89	1031	686.08	3.793×10^{-73}
Salt	3	1.99	1028	684.09	0.58
Protein:pH	63	586.48	965	97.61	5.986×10^{-86}
Protein:salt	63	19.15	902	78.46	1
pH:salt	9	12.99	893	65.47	1.60×10^{-1}
Protein:pH:salt	189	59.79	704	5.67	1

Table A1.2 - Analysis of deviance table for the statistical analysis of DSF of all *E. coli* DHDPS variants.

A1.2 CIRCULAR DICHROISM SPECTROSCOPY

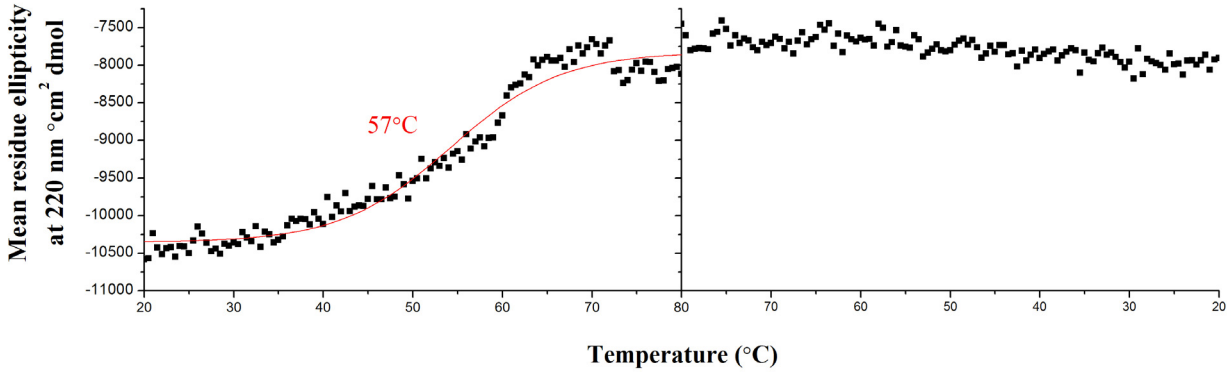


Figure A1.2 – Circular dichroism spectroscopy of wild-type *E. coli* DHDPS (0.05 mg/mL) in 20 mM phosphate buffer (pH 7, 40 mM NaCl). Mean residue ellipticity at 220 nm vs. temperature indicating melting and attempted refold of the protein. The red line corresponds to a sigmoid (Boltzmann) fit and the temperature shown was calculated by identifying the midpoint of the transition.

The thermal denaturation temperatures at pH 2 and pH 11 were investigated. These yielded thermal denaturation temperatures significantly different to that of wild type protein, melting at 55 ºC and 52 ºC respectively and no significant refolding events. The presence of 40 mM salt at both pH 2 and pH 11 lead to elevated voltage resulting in noise that compromised the results of the thermal denaturation assay, thus the thermal denaturation assays were performed without salt. This significantly affected the results of the assay; in pH 2 and pH 11 buffer (with no salt) proteins appear exist in a partially unfolded state, particularly at pH 2. This was corroborated by published data which found that many proteins, under denaturing conditions, can retain residual structure [1]. The presence of 40 mM NaCl appears to enable *E. coli* DHDPS to retain a more native-like secondary structure which is observed throughout the range of pH tested (section 2.2).

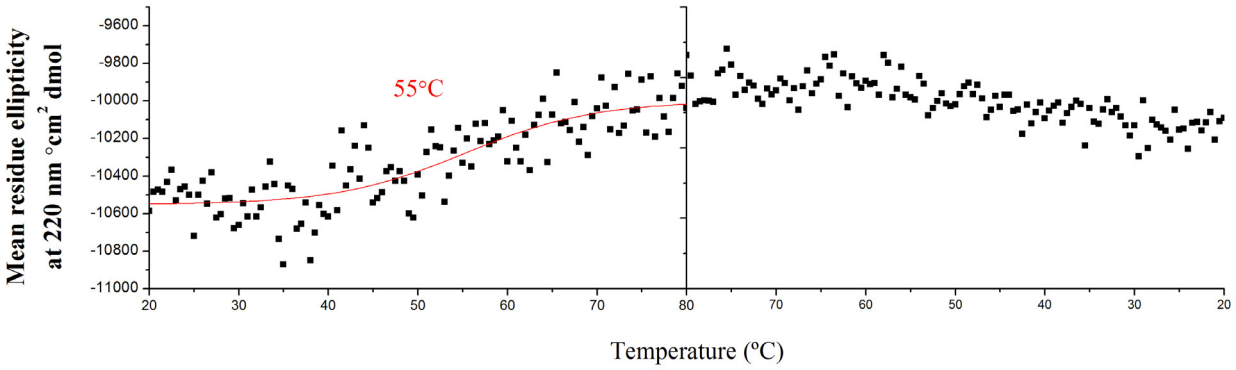


Figure A1.3 – Circular dichroism spectroscopy of wild-type *E. coli* DHDPS (0.05 mg/mL) in 20 mM phosphate buffer (pH 2). Mean residue ellipticity at 220 nm vs. temperature indicating melting and attempted refold of the protein. The red line corresponds to a sigmoid (Boltzmann) fit and the temperature shown was calculated by identifying the midpoint of the transition.

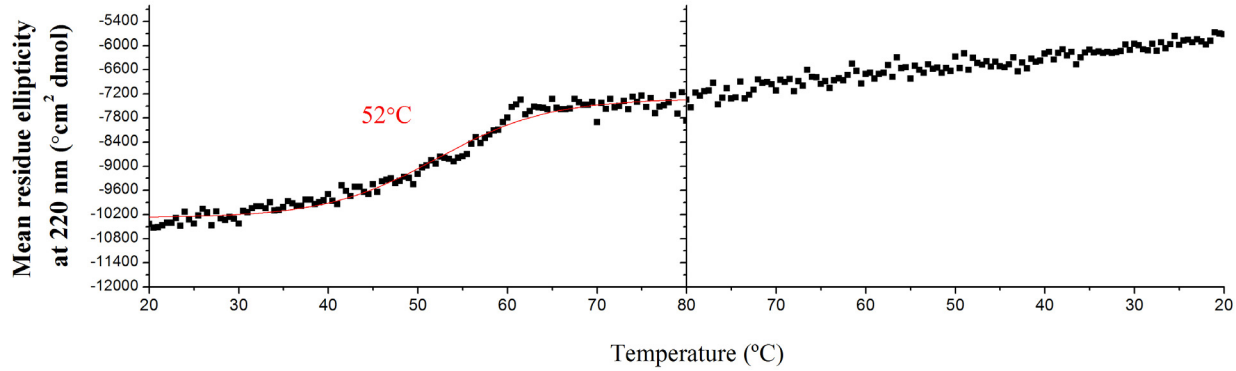


Figure A1.4 – Circular dichroism spectroscopy of wild-type *E. coli* DHDPS (0.05 mg/mL) in 20 mM phosphate buffer (pH 11). Mean residue ellipticity at 220 nm vs. temperature indicating melting and attempted refold of the protein. The red line corresponds to a sigmoid (Boltzmann) fit and the temperature shown was calculated by identifying the midpoint of the transition.

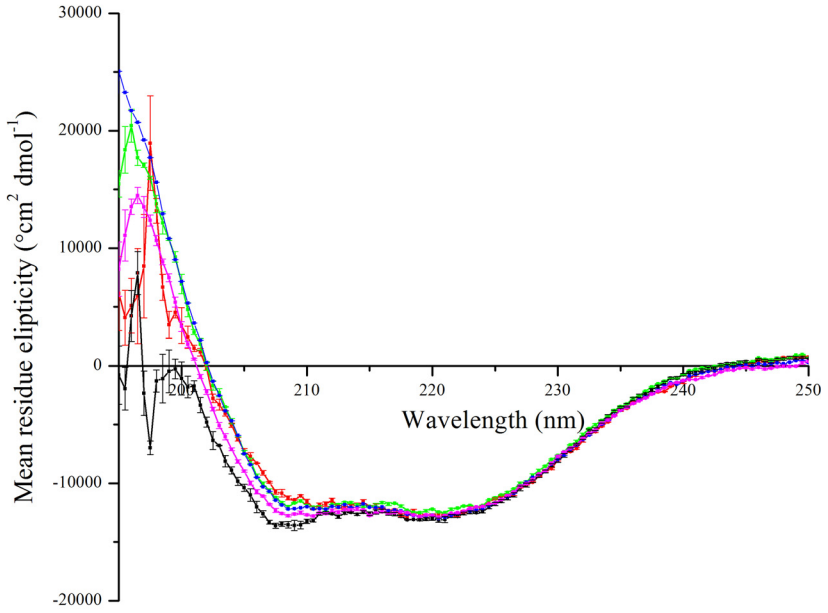


Figure A1.5 – Circular dichroism spectroscopy of *E. coli* DHDPS in 20 mM phosphate buffer, 40 mM NaCl at pH 2 (red), pH 4.5 (green), pH 7 (blue), pH 9.5 (pink) and pH 11 (black) at 0.2 mg/mL. Mean residue ellipticity is plotted against wavelength.

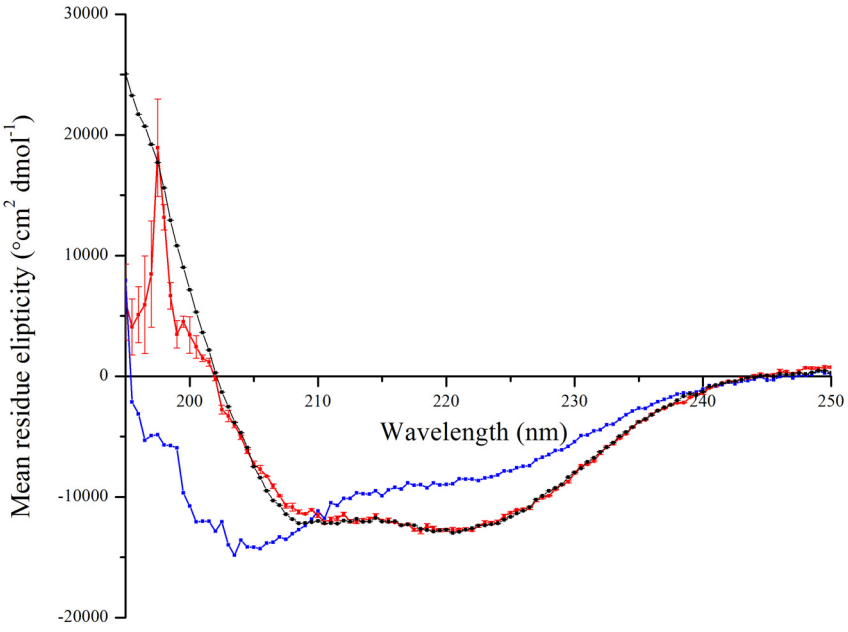


Figure A1.6 – Circular dichroism spectroscopy of *E. coli* DHDPS in 20 mM phosphate buffer at pH 7 (black), pH 2, 0 mM NaCl (blue), and pH 2, 40 mM NaCl (red) at 0.2 mg/mL. Mean residue ellipticity is plotted against wavelength.

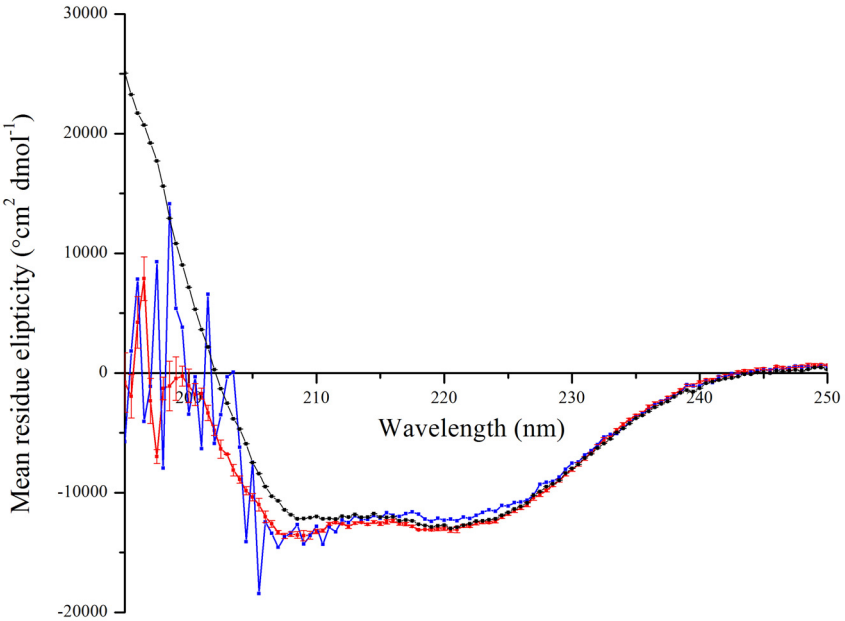


Figure A1.7 – Circular dichroism spectroscopy of *E. coli* DHDPS in 20 mM phosphate buffer at pH 7 (black), pH 11, 0 mM NaCl (blue), and pH 11, 40 mM NaCl (red) at 0.2 mg/mL. Mean residue ellipticity is plotted against wavelength.

A1.3 AMORPHOUS AGGREGATION

pH	Salt concentration (mM)	*Agg*½ (minutes) ± SEM	Maximum absorbance (A.U.) ± SEM
2	0		0.13 ± 0.02
	20		0.08 ± 0.02
	40		0.12 ± 0.01
	100		0.14 ± 0.03
	500		1.13 ± 0.18
4.5	0		1.19 ± 0.07
	20		1.22 ± 0.04
	40		1.24 ± 0.02
	100		1.14 ± 0.05
	500		1.13 ± 0.05
7	0	46.8 ± 0.6	1.08 ± 0.04
	20	47.9 ± 0.2	1.02 ± 0.03
	40	47.8 ± 0.2	1.01 ± 0.02
	100	45.6 ± 0.4	0.99 ± 0.02
	500	35.0 ± 0.5	0.80 ± 0.02
9.5	0	70.0 ± 0.9	0.48 ± 0.02
	20	62.8 ± 0.5	0.59 ± 0.02
	40	60.6 ± 1.2	0.64 ± 0.03

	100	56.7 ± 1.6	0.77 ± 0.05
	500	37.5 ± 0.9	0.80 ± 0.02
11	0		0.11 ± 0.01
	20		0.16 ± 0.01
	40		0.16 ± 0.01
	100		0.14 ± 0.01
	500		0.14 ± 0.01

Table A1.3 –The aggregation half life ($Agg_{½}$) of wild-type *E. coli* DHDPS as determined by the calculation of the $t_{½}$ of a sigmoid function fitted to the data in figure 2.18. Errors were calculated as per the CD data (figure 2.17). The maximum absorbance was the maximum reached over the initial 90 minutes of the assay. The errors were calculated from the mean of the six replicates.

	Degrees of freedom	Sum square	Mean square	F value	Pr (>F)
pH	1	2497.2	2497.2	638.3	$< 2.2 \times 10^{-16}$
salt	4	3610	902.5	230.7	$< 2.2 \times 10^{-16}$
pH:salt	4	661.5	165.4	42.3	1.85×10^{-15}
residuals	50	195.6	3.9		

Table A1.4 - Analysis of deviance table for the statistical analysis of the $Agg_{½}$ of wild-type DHDPS.

	Degrees of freedom	Sum of squares	Mean squares	F value	Pr (>F)
pH	4	19.40	4.85	52.80	3.0×10^{-6}
Salt	4	3.12	0.80	8.73	$< 2.2 \times 10^{-16}$
pH:salt	16	5.63	0.35	3.83	8.9×10^{-6}
Residuals	125	11.48	0.09		

Table A1.5 - Analysis of deviance table for the statistical analysis of the maximum absorbance of wild-type DHDPS.

A1.4 β-SHEET-SPECIFIC AGGREGATION

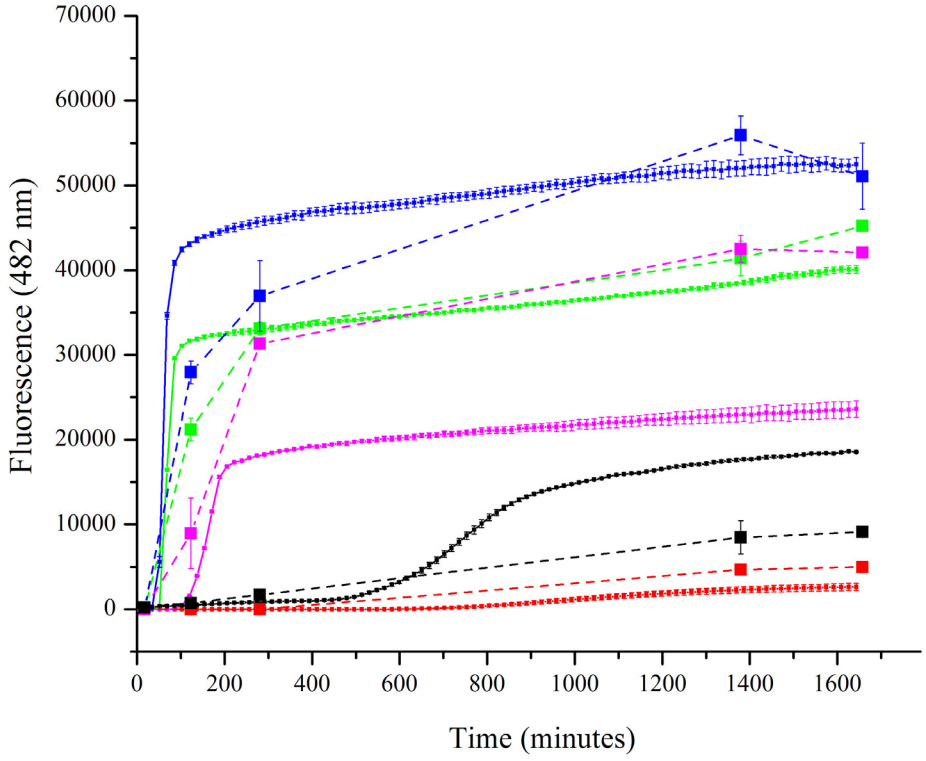


Fig A1.8 – The thioflavin T induced fluorescence of insulin in 0 mM NaCl (red), 20 mM NaCl (green), 40 mM NaCl (blue), 100 mM NaCl (pink) and 500 mM NaCl (black). The data represented by the small symbols are time course measurements in which the ThT was added at time 0. The large data points represent data obtained through the addition of ThT at time points during incubation. Note that the dashed line is present only to guide the eye. The variability between the time course data and the data obtained from the addition of ThT at intervals revealed that salt concentration significantly affects both fibril formation and ThT binding. The data plotted is the mean of three replicates and the error bars represent the SEM.

pH	Salt concentration (mM)	β-$agg_{½}$ (minutes) ± SEM	Maximum absorbance (AU) ± SEM
2	0		-66 ± 19
	40		-49 ± 19
	500	82.8 ± 8.4	795 ± 51
4.5	0	65.0 ± 1.6	2256 ± 74
	40	60.4 ± 4.6	1953 ± 434
	500	82.0 ± 2.6	4407 ± 229
7	0	77.9 ± 1.6	4701 ± 273
	40	84.0 ± 2.3	4861 ± 322
	500	100.2 ± 2.3	4513 ± 272
9.5	0		859 ± 61
	40		815 ± 11
	500		1348 ± 58
11	0		-12 ± 16
	40		12 ± 15
	500		393 ± 14

Table A1.6 – The β-sheet specific aggregation half life (β-$agg_{½}$) of wild-type DHDPS as determined by the calculation of the $t_{½}$ of a sigmoid (Boltzmann) function fitted to the ThT fluorescence data (chapter 2, figure 2.21). Errors were calculated as per the CD data (chapter 2, figure 2.17). The maximum fluorescence was the maximum reached over the initial 300 minutes of the assay. The errors were calculated from the mean of the six replicates.

A1.5 REFERENCES

1 Hammarström, M. and Carlsson, J. (2000) Is the unfolded state the Rosetta stone of the protein folding problem? Biochemical and Biophysical Research Communication **276**, 393 - 398

APPENDIX 2

SUPPLEMENTARY INFORMATION FOR CHAPTER 3

A2.1 ALGORITHMIC PREDICTION

	Wild-type DHDPS	pET M11 DHDPS	pET 151/D-TOPO DHDPS	cleaved pET 151/D-TOPO DHDPS
Overall hydrophobicity at pH 7	95.38	105.83	99.03	95.11
Overall charge at pH 7	-5	-7	-7	-6
Overall β-sheet propensity	1090.06	1186.27	1195.4	1105.69
Overall α-helical propensity	1133.27	1234.35	1250.29	1151.29
Number of hydrophobic/hydrophilic patterns	2	2	2	2
Absolute amyloid aggregation rate	-4.07	-4.25	-4.26	-4.15
Intrinsic aggregation propensity	-6.12	-6.74	-6.87	-6.29
Average ± SD	-5.46 ± 1.45	-5.72 ± 1.52	-5.81 ± 1.53	-5.52 ± 1.47
Zagg score	-0.46	-0.67	-0.69	-0.52

Table A2.1 – The results of the Zyggregator algorithm for wild-type DHDPS, pET M11 DHDPS, pET 155/D-TOPO DHDPS and cleaved pET 151/D-TOPO DHDPS.

A2.2 DIFFERENTIAL SCANNING FLUORIMETRY

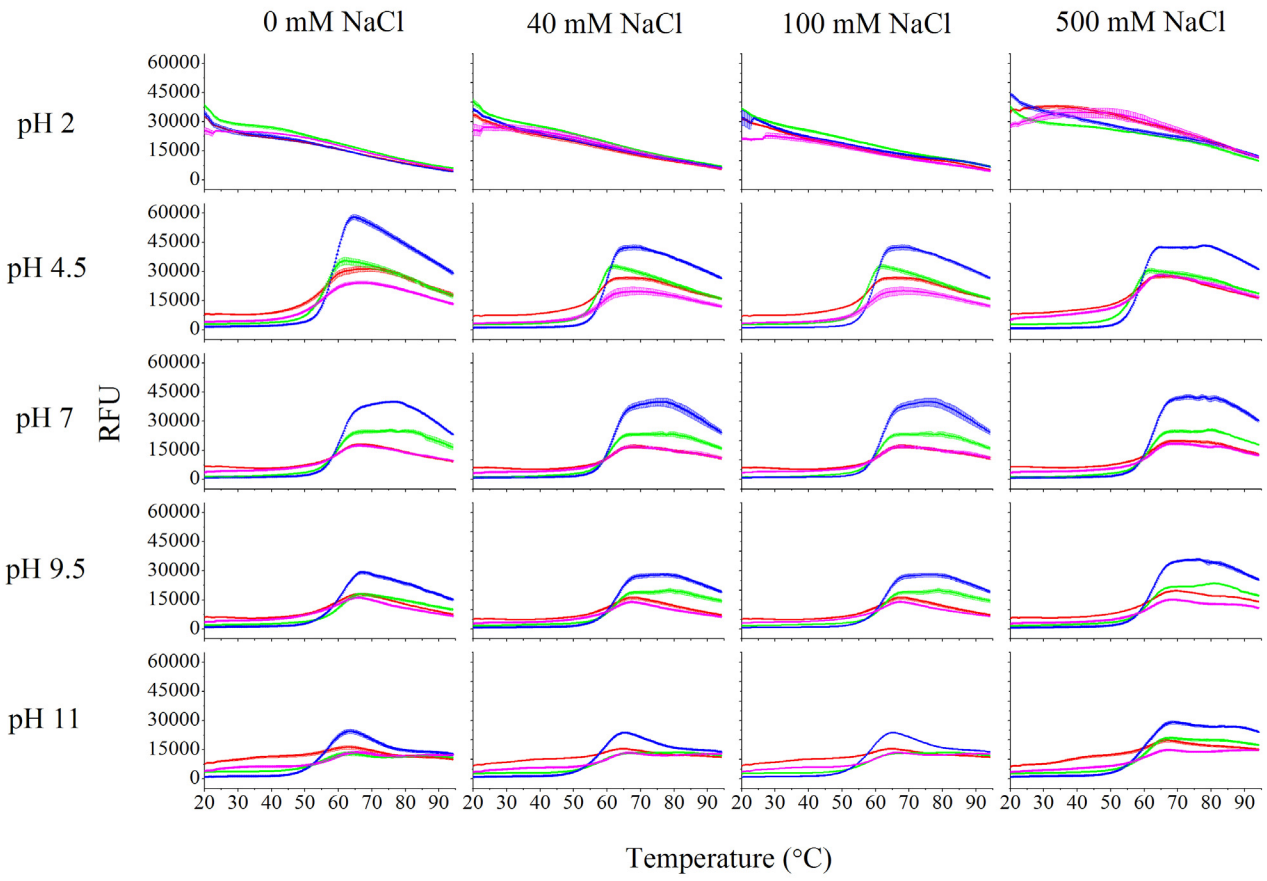


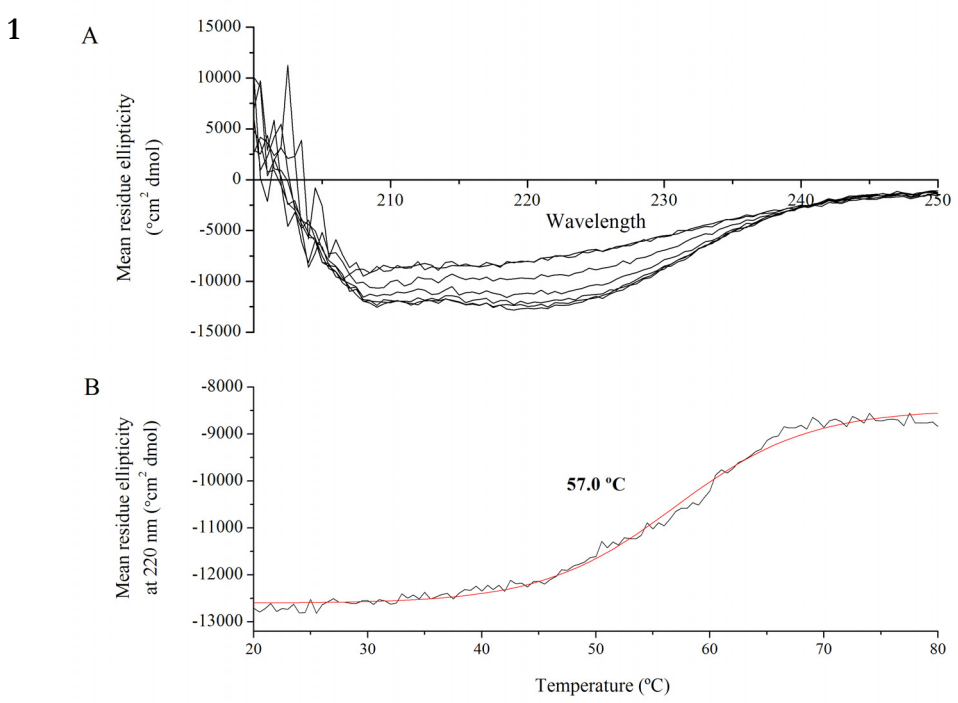
Figure A2.1 – The thermal denaturation of wild-type DHDPS (red), pET M11 DHDPS (green), pET 151/D-TOPO DHDPS (blue) and cleaved pET 151/D-TOPO DHDPS (pink) at 0.5 mg/mL in 100 mM phosphate buffer at pH 2, pH 4.5, pH 7, pH 9.5 and pH 11 containing 0 mM NaCl, 40 mM NaCl, 100 mM NaCl and 500 mM NaCl as monitored by SYPRO orange fluorescence. The data plotted are the increase in fluorescence, as monitored by the BioRad IQ5.

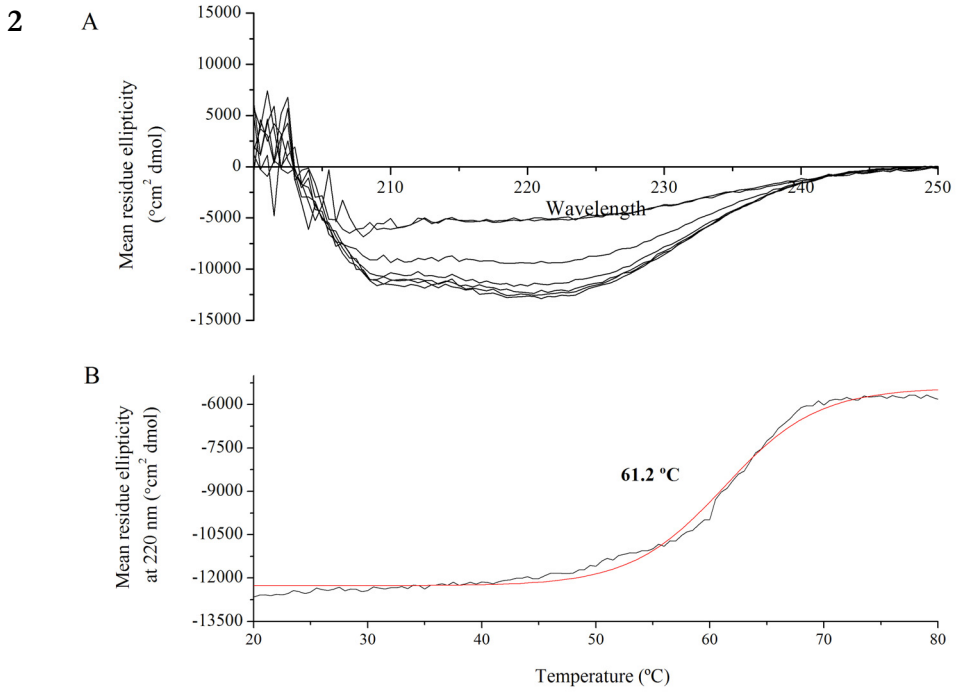
Variant	pH	Salt concentration	Mean T_m (°C)
Wild-type DHDPS	4.5	0 mM	57.8 ± 0.5
		40 mM	57.7 ± 0.2
		100 mM	57.9 ± 0.2
		500 mM	57.7 ± 0.2
	7	0 mM	60.5 ± 0.3
		40 mM	60.9 ± 0.3
		100 mM	61.5 ± 0.1
		500 mM	61.8 ± 0.2
	9.5	0 mM	60.8 ± 1.0
		40 mM	61.9 ± 0.4
		100 mM	61.9 ± 0.3
		500 mM	62.2 ± 0.1
	11	0 mM	57.2 ± 0.3
		40 mM	57.3 ± 0.4
		100 mM	59.0 ± 0.0
		500 mM	59.2 ± 0.1
pET M11 polyhistidine-tagged DHDPS	4.5	0 mM	56.9 ± 0.5
		40 mM	56.5 ± 0.5
		100 mM	57.0 ± 0.5
		500 mM	57.5 ± 0.2
	7	0 mM	59.8 ± 0.2
		40 mM	60.2 ± 0.0
		100 mM	60.2 ± 0.0
		500 mM	60.2 ± 0.0
	9.5	0 mM	61.4 ± 0.0
		40 mM	61.6 ± 0.1
		100 mM	61.8 ± 0.1
		500 mM	61.9 ± 0.1
	11	0 mM	55.9 ± 0.1
		40 mM	55.9 ± 0.1
		100 mM	56.0 ± 0.0
		500 mM	58.0 ± 0.5
pET 151/D-TOPO DHDPS	4.5	0 mM	57.7 ± 0.1
		40 mM	59.6 ± 0.0
		100 mM	59.6 ± 0.0
		500 mM	59.7 ± 0.1
	7	0 mM	62.0 ± 0.0
		40 mM	62.0 ± 0.0
		100 mM	62.0 ± 0.0
		500 mM	62.3 ± 0.2
	9.5	0 mM	61.4 ± 0.3
		40 mM	61.9 ± 0.4
		100 mM	62.1 ± 0.1
		500 mM	62.2 ± 0.1
	11	0 mM	57.5 ± 0.0
		40 mM	57.5 ± 0.0

		100 mM	57.5 ± 0.0
		500 mM	57.5 ± 0.0
Cleaved pET 151/D-TOPO DHDPS	4.5	0 mM	55.9 ± 0.4
		40 mM	57.0 ± 0.2
		100 mM	57.2 ± 0.2
		500 mM	57.9 ± 0.2
	7	0 mM	59.9 ± 0.0
		40 mM	60.8 ± 0.3
		100 mM	61.3 ± 0.2
		500 mM	61.7 ± 0.0
	9.5	0 mM	60.2 ± 0.4
		40 mM	61.8 ± 0.1
		100 mM	62.1 ± 0.2
		500 mM	62.2 ± 0.1
	11	0 mM	59.8 ± 0.1
		40 mM	59.8 ± 0.1
		100 mM	60.3 ± 0.4
		500 mM	60.3 ± 0.4

Table A2.2 – The thermal denaturation temperatures of wild-type DHDPS, pET M11 DHDPS, pET 151/D-TOPO DHDPS and cleaved pET 151/D-TOPO DHDPS in 100 mM phosphate buffer as monitored by SYPRO orange fluorescence and as assessed from the derivation of the fluorescence data.

A2.3 CIRCULAR DICHROISM SPECTROSCOPY





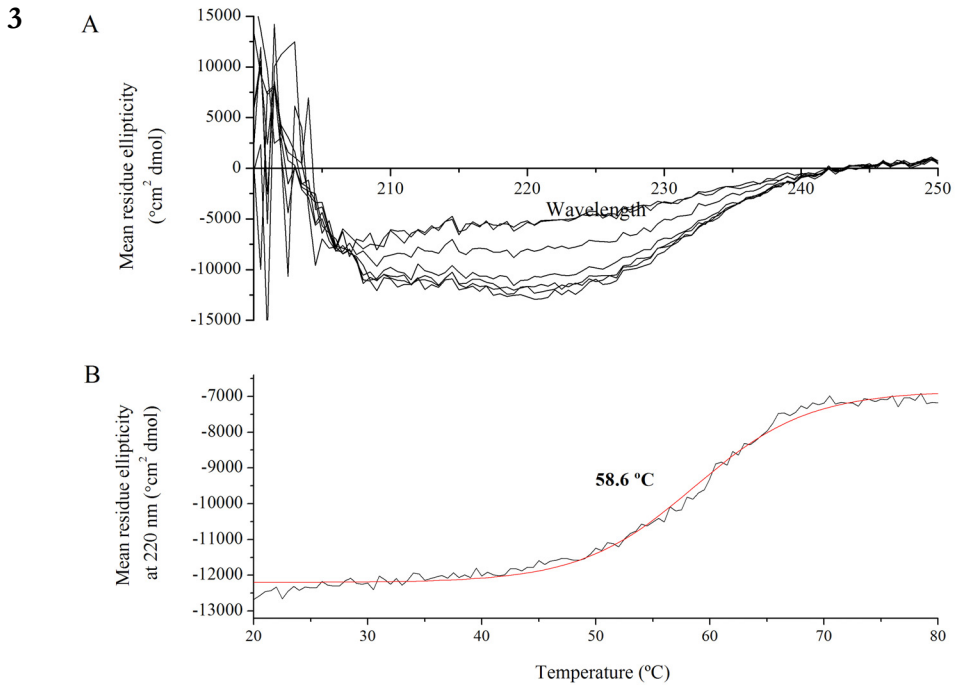


Figure A2.2 –Circular dichroism spectroscopy of **(1)** pET M11 DHDPS, **(2)** pET 151/D-TOPO DHDPS and **(3)** cleaved pET 151/D-TOPO DHDPS. **A)** Wavelength scans plotting mean residue ellipticity vs. wavelength at 20°C, 30°C, 40°C, 50°C, 60°C, 70°C and 80°C. **B)** Mean residue ellipticity [θ] vs. temperature indicating melting of the protein. The red line corresponds to a sigmoid (Boltzmann) fit and the temperature shown is the midpoint of that fit.

A2.4 AMORPHOUS AGGREGATION

Variant	pH	Salt concentration (mM)	$Agg_{½}$ (minutes) ± SEM	Maximum absorbance (AU) ± SEM
Wild-type DHDPS	2	0		0.13 ± 0.02
		20		0.08 ± 0.02
		40		0.12 ± 0.01
		100		0.14 ± 0.03
		500		1.13 ± 0.18
	4.5	0		1.19 ± 0.07
		20		1.22 ± 0.04
		40		1.24 ± 0.02
		100		1.14 ± 0.05
		500		1.13 ± 0.05
	7	0	46.8 ± 0.6	1.08 ± 0.04
		20	47.9 ± 0.2	1.02 ± 0.03
		40	47.8 ± 0.2	1.01 ± 0.02

		100	45.6 ± 0.4	0.99 ± 0.02
		500	35.0 ± 0.5	0.80 ± 0.02
	9.5	0	70.0 ± 0.9	0.48 ± 0.02
		20	62.8 ± 0.5	0.59 ± 0.02
		40	60.6 ± 1.2	0.64 ± 0.03
		100	56.7 ± 1.6	0.77 ± 0.05
		500	37.5 ± 0.9	0.80 ± 0.02
	11	0		0.11 ± 0.01
		20		0.16 ± 0.01
		40		0.16 ± 0.01
		100		0.14 ± 0.01
		500		0.14 ± 0.01
pET M11 DHDPS	2	0		0.11 ± 0.01
		20		0.09 ± 0.02
		40		0.11 ± 0.01
		100		0.07 ± 0.01
		500	13.3 ± 5.4	0.64 ± 0.04
	4.5	0		1.32 ± 0.05
		20		1.38 ± 0.04
		40		1.31 ± 0.05
		100		1.20 ± 0.04
		500	35.6 ± 0.4	1.10 ± 0.03
	7	0	36.0 ± 0.7	0.88 ± 0.03
		20	39.6 ± 0.3	0.87 ± 0.02
		40	40.6 ± 0.2	0.81 ± 0.03
		100	40.8 ± 0.3	0.76 ± 0.03
		500	37.1 ± 0.3	0.78 ± 0.03
	9.5	0	46.3 ± 1.0	0.57 ± 0.03
		20	48.0 ± 0.7	0.50 ± 0.02
		40	42.3 ± 1.1	0.52 ± 0.04
		100	41.9 ± 0.5	0.53 ± 0.02
		500	35.6 ± 0.4	0.68 ± 0.01
	11	0		0.11 ± 0.01
		20		0.09 ± 0.01
		40		0.11 ± 0.02
		100		0.07 ± 0.01
		500	56.3 ± 1.9	0.64 ± 0.07
pET 151/D-TOPO DHDPS	2	0		0.13 ±0.02
		20		0.17 ± 0.01
		40		0.14 ± 0.01
		100		0.14 ± 0.01
		500	9.1 ± 0.5	0.42 ± 0.05
	4.5	0		1.43 ± 0.06
		20		1.37 ± 0.05
		40	29.4 ± 0.8	1.36 ± 0.06
		100	32.3 ± 0.9	1.44 ± 0.03
		500	36.3 ± 0.9	1.10 ± 0.07
	7	0	28.6 ± 0.8	0.76 ± 0.02
		20	33.8 ± 0.5	0.72 ± 0.03
		40	34.2 ± 0.4	0.73 ± 0.03

		100	35.0 ± 0.9	0.65 ± 0.02
		500	31.4 ± 0.4	0.68 ± 0.03
	9.5	0	33.0 ± 0.9	0.34 ± 0.02
		20	34.1 ± 0.2	0.51 ± 0.10
		40	33.9 ± 0.5	0.60 ± 0.09
		100	31.3 ± 0.6	0.42 ± 0.05
		500	30.8 ± 0.3	0.53 ± 0.02
	11	0		0.13 ± 0.01
		20		0.17 ± 0.06
		40		0.14 ± 0.01
		100		0.14 ± 0.04
		500	35.4 ± 1.5	0.37 ± 0.03
Cleaved pET 151/D-TOPO DHDPS	2	0		0.10 ± 0.01
		20		0.09 ± 0.02
		40		0.11 ± 0.02
		100		0.15 ± 0.02
		500		0.97 ± 0.04
	4.5	0	27.6 ± 0.9	1.11 ± 0.01
		20	30.3 ± 0.3	1.09 ± 0.00
		40	31.5 ± 0.4	1.12 ± 0.00
		100	32.3 ± 0.3	1.04 ± 0.03
		500	33.8 ± 0.6	0.98 ± 0.02
	7	0	33.3 ± 0.9	0.60 ± 0.01
		20	35.3 ± 0.4	0.57 ± 0.01
		40	37.4 ± 0.6	0.56 ± 0.01
		100	37.0 ± 0.3	0.55 ± 0.00
		500	32.9 ± 0.4	0.52 ± 0.00
	9.5	0	42.5 ± 0.9	0.34 ± 0.02
		20	40.4 ± 0.1	0.22 ± 0.01
		40	42.0 ± 0.2	0.40 ± 0.02
		100	42.1 ± 0.2	0.28 ± 0.01
		500	38.7 ± 0.3	0.40 ± 0.01
	11	0		0.09 ± 0.01
		20		0.05 ± 0.02
		40		0.11 ± 0.01
		100		0.07 ± 0.03
		500	37.3 ± 0.4	0.32 ± 0.01

Table A2.3 – The aggregation half life ($Agg_{½}$) and maxima of wild-type DHDPS, pET M11 DHDPS, pET 151/D-TOPO DHDPS and cleaved pET 151/D-TOPO DHDPS determined by the calculation of the $t_{½}$ of a sigmoid function fitted to the data in chapter 3, figure 3.9. Errors were calculated as per the CD data (figure 2.17). The maximum absorbance was the maximum reached over the initial 90 minutes of the assay. The errors were calculated from the mean of the six replicates.

	Degrees of freedom	Sum of squares	Mean squares	F value	Pr (>F)
Protein	3	13766.5	4588.8	1760.4	< 2.2 x 10-16
pH	1	1073.6	1073.6	411.9	< 2.2 x 10-16
Salt	4	1972.5	493.1	189.2	< 2.2 x 10-16
Protein:pH	3	1660.7	553.6	212.4	< 2.2 x 10-16
Protein:salt	12	2225.6	185.5	71.1	< 2.2 x 10-16
pH:salt	4	857.3	214.3	82.2	< 2.2 x 10-16
protein:pH:salt	12	315.6	26.3	10.1	< 2.2 x 10-15
Residuals	200	521.3	2.6		

Table A2.4 - Analysis of deviance table for the statistical analysis of the $Agg_{½}$ of *E. coli* DHDPS and its polyhistidine-tagged variants.

	Degrees of freedom	Sum of squares	Mean squares	F value	Pr (>F)
Protein	3	2.65	0.88	8.75	1.2×10^{-5}
pH	4	70.25	17.56	174.25	$< 2.2 \times 10^{-16}$
Salt	4	14.37	3.59	35.64	$< 2.2 \times 10^{-16}$
Protein:pH	12	2.37	0.20	1.96	2.6×10^{-2}
Protein:salt	12	0.25	0.02	0.20	1.00
pH:salt	16	16.07	1.01	9.97	$< 2.2 \times 10^{-16}$
protein:pH:salt	48	1.37	0.03	0.28	1.00
Residuals	500	50.40	0.10		

Table A2.5 - Analysis of deviance table for the statistical analysis of the maximum absorbance of *E. coli* DHDPS and its polyhistidine-tagged variants.

A2.5 β-SHEET SPECIFIC AGGREGATION

Variant	pH	Salt concentration (mM)	β-$agg_{½}$ (minutes) ± SEM	Maximum fluorescence (RFU) ± SEM
Wild-type DHDPS	2	0		-66 ± 19
		40		-49 ± 19
		500	82.8 ± 8.4	795 ± 51
	4.5	0	65.0 ± 1.6	2256 ± 74
		40	60.4 ± 4.6	1953 ± 434
		500	82.0 ± 2.6	4407 ± 229
	7	0	77.9 ± 1.6	4701 ± 273
		40	84.0 ± 2.3	4861 ± 322
		500	100.2 ± 2.3	4513 ± 272
	9.5	0		859 ± 61
		40		815 ± 11
		500		1348 ± 58

	11	0		-12 ± 16
		40		12 ± 15
		500		393 ± 14
pET M11 DHDPS	2	0		62 ± 13
		40		69 ± 14
		500		538 ± 70
	4.5	0	66.78 ± 1.92	4428 ± 251
		40	70.31 ± 1.37	4290 ± 368
		500	73.57 ± 2.55	4702 ± 346
	7	0	65.59 ± 0.53	9963 ± 25
		40	72.16 ± 0.26	9120 ± 561
		500	75.86 ± 0.49	6464 ± 351
	9.5	0		522 ± 25
		40		553 ± 25
		500		1728 ± 94
	11	0		-52 ± 31
		40		-43 ± 20
		500		-22 ± 16
pET 151/D-TOPO DHDPS	2	0		-39 ± 30
		40		-62 ± 29
		500		495 ± 70
	4.5	0	63.70 ± 1.01	2750 ± 129
		40	67.89 ± 0.36	3169 ± 156
		500	75.39 ± 1.16	4201 ± 190
	7	0	71.99 ± 0.89	6812 ± 203
		40	77.51 ± 0.64	6396 ± 362
		500	77.15 ± 1.36	4991 ± 183
	9.5	0		269 ± 26
		40		276 ± 24
		500		1391 ± 61
	11	0		-82 ± 19
		40		-19 ± 32
		500		-31 ± 42
Cleaved pET 151/D-TOPO DHDPS	2	0		-118 ± 34
		40		-224 ± 10
		500		-120 ± 17
	4.5	0	67.33 ± 1.64	678 ± 25
		40	70.61 ± 1.34	700 ± 19
		500	74.41 ± 2.51	924 ± 42
	7	0	65.66 ± 0.53	2072 ± 61
		40	72.25 ± 0.26	1947 ± 31
		500	75.53 ± 0.54	1585 ± 31
	9.5	0		28 ± 8
		40		-41 ±8
		500		288 ± 20
	11	0		-96 ± 5
		40		-62 ± 8
		500		45 ± 6

Table A2.6 –The β-sheet specific aggregation half life (β-$agg_{½}$) of wild-type DHDPS and its polyhistidine-tagged variants as determined by the calculation of the $t_{½}$ of a sigmoid (Boltzmann) function fitted to the ThT fluorescence data (chapter 3, figure 3.11). Errors were calculated as per the CD data (chapter 2, figure 2.17). The maximum fluorescence was the maximum reached over the initial 300 minutes of the assay. The errors were calculated from the mean of the six replicates.

	Degrees of freedom	Sum of squares	Mean squares	F value	Pr (>F)
Protein	3	1346	449	28	1.6×10^{-12}
Salt	2	3151	1576	87	$< 2.2 \times 10^{-16}$
pH:salt	1	1534	1534	85	1.4×10^{-15}
Protein:salt	6	1141	190	10	2.5×10^{-9}
Protein:pH	3	1845	615	34	6.0×10^{-16}
salt:pH	2	135	67	4	2.7×10^{-2}
Protein:salt:pH	6	183	30	2	1.3×10^{-1}
Residuals	120	2176	18		

Table A2.7 - Analysis of deviance table for the statistical analysis of the β-$agg_{½}$ of *E. coli* DHDPS and its polyhistidine-tagged variants. Note that pH and salt were included in the model because they had an effect on the fluorescence; however, this was not relevant due to their affect on ThT binding. The main effect of protein was the independent effect after controlling between the differences across pH and salt concentrations.

	Degrees of freedom	Sum of squares	Mean squares	F value	Pr (>F)
Protein	3	211978686	70659562	385.7	$< 2.2 \times 10^{-16}$
pH	4	1539288769	384822192	2100.8	$< 2.2 \times 10^{-16}$
Salt	2	2295701	1147851	6.3	2.2×10^{-3}
Protein:pH	12	329767842	27480653	150	$< 2.2 \times 10^{-16}$
Protein:salt	6	11658257	1943043	10.6	1.1×10^{-10}
pH:salt	8	45644041	5705505	31.4	$< 2.2 \times 10^{-16}$
Protein:pH:salt	24	18046991	751958	4.1	2.8×10^{-9}
Residuals	300	54953600	183179		

Table A2.8 - Analysis of deviance table for the statistical analysis of the maximum fluorescence of *E. coli* DHDPS and its polyhistidine-tagged variants. As discussed above, pH and salt were included in the model because they had an effect on the fluorescence; however, this was not relevant due to their affect on ThT binding.

APPENDIX 3

SUPPLEMENTARY INFORMATION FOR CHAPTER 4

A3.1 ALGORITHMIC PREDICTION

	Wild-type DHDPS	DHDPS Q90L	pET M11 DHDPS Q90L	pET 151/D-TOPO DHDPS Q90L	cleaved pET 151/D-TOPO DHDPS Q90L
Overall hydrophobicity at pH 7	95.38	92.26	102.71	95.91	91.99
Overall charge at pH 7	-5	-5	-7	-7	-6
Overall β-sheet propensity	1090.06	1090.25	1186.46	1195.59	1105.88
Overall α-helical propensity	1133.27	1133.52	1234.6	1250.53	1151.54
Number of hydrophobic/hydrophilic patterns	2	6	6	6	6
Absolute amyloid aggregation rate	-4.07	-2.47	-2.66	-2.67	-2.55
Intrinsic aggregation propensity	-6.12	-4.51	-5.13	-5.26	-4.68
Average ± SD	-5.46 ± 1.45	-5.46 ± 1.45	-5.73 ± 1.52	-5.81 ± 1.53	-5.51 ± 1.47
Zagg score	-0.458258	0.651956	0.398088	0.359207	0.567079

Table A3.1 – The results of the Zyggregator algorithm for wild-type DHDPS, DHDPS Q90L, pET M11 DHDPS Q90L, pET 151/D-TOPO DHDPS Q90L and cleaved pET 151/D-TOPO DHDPS Q90L.

	Wild-type DHDPS	DHDPS A207V	pET M11 DHDPS A207V	pET 151/D-TOPO DHDPS A207V	cleaved pET 151/D-TOPO DHDPS A207V
Overall hydrophobicity at pH 7	95.38	94.47	104.92	98.12	94.2
Overall charge at pH 7	-5	-5	-7	-7	-6
Overall β-sheet propensity	1090.06	1090.47	1186.68	1195.81	1106.1
Overall α-helical propensity	1133.27	1133.29	1234.36	1250.3	1151.31
Number of hydrophobic/hydrophilic patterns	2	5	5	5	5
Absolute amyloid aggregation rate	-4.07	-2.87	-3.06	-3.07	-2.95
Intrinsic aggregation propensity	-6.12	-4.92	-5.53	-5.66	-5.09
Average ± SD	-5.46 ± 1.45	-5.46 ± 1.46	-5.73 ± 1.51	-5.80 ± 1.53	-5.52 ± 1.47
Zagg score	-0.458258	0.372337	0.134293	0.0920158	0.293241

Table A3.2 – The results of the Zyggregator algorithm for wild-type DHDPS, DHDPS A207V, pET M11 DHDPS A207V, pET 151/D-TOPO DHDPS A207V and cleaved pET 151/D-TOPO DHDPS A207V.

A3.2 DIFFERENTIAL SCANNING FLUORIMETRY

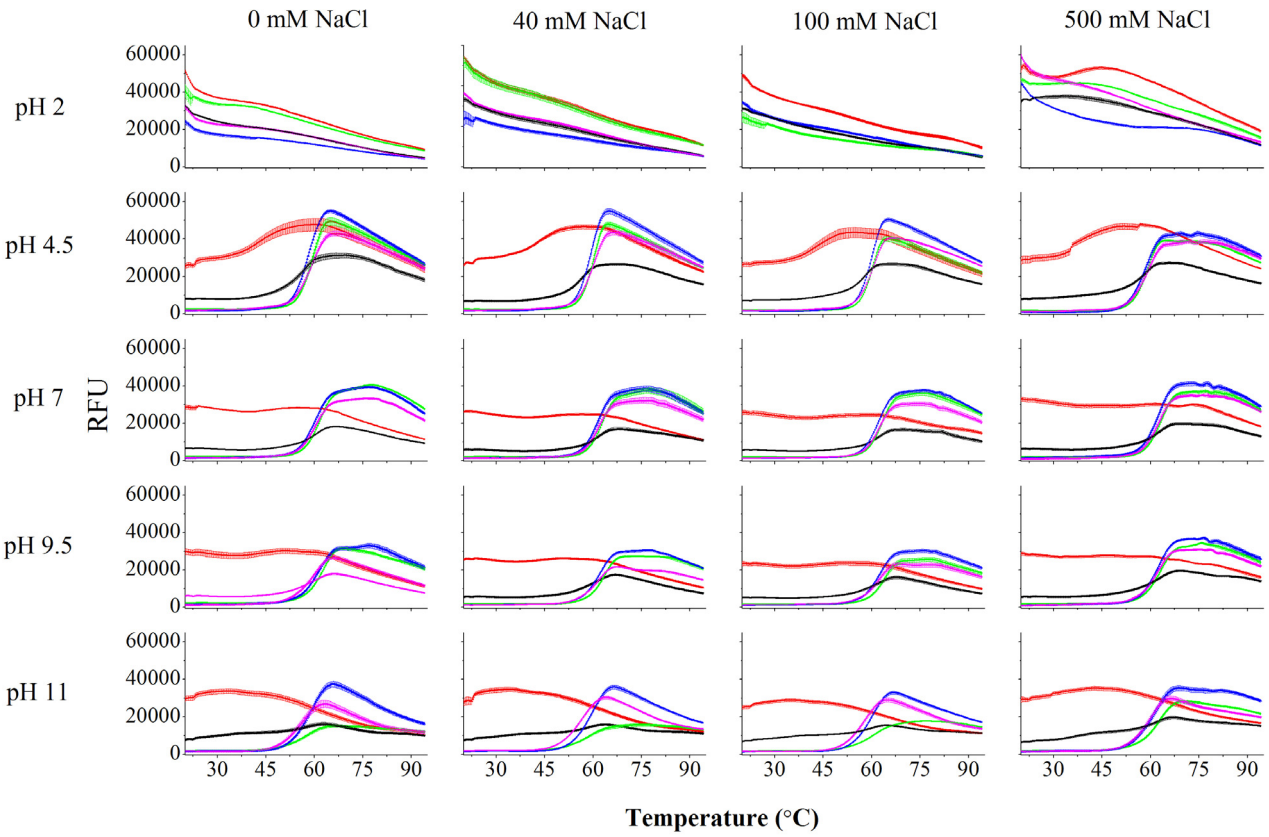


Figure A3.1 – The thermal denaturation of wild-type DHDPS (black), DHDPS Q90L (red) pET M11 DHDPS Q90L (green), pET 151/D-TOPO DHDPS Q90L (blue) and cleaved pET 151/D-TOPO DHDPS Q90L (pink) in 100 mM phosphate buffer at pH 2, pH 4.5, pH 7, pH 9.5 and pH 11 containing 0 mM NaCl, 40 mM NaCl, 100 mM NaCl and 500 mM NaCl as monitored by SYPRO Orange fluorescence. The data plotted are the mean increase in fluorescence, as monitored by the BioRad IQ5 ± SEM.

Variant	pH	Salt concentration	Mean thermal denaturation temperature 1 (°C) ± SEM	Mean thermal denaturation temperature 2 (°C) ± SEM
pET M11 DHDPS Q90L	4.5	0 mM	58.9 ± 0.2	
		40 mM	60.2 ± 0.0	
		100 mM	60.2 ± 0.0	
		500 mM	60.0 ± 0.1	
	7	0 mM	63.0 ± 0.1	
		40 mM	63.1 ± 0.1	
		100 mM	63.0 ± 0.1	
		500 mM	63.1 ± 0.1	

	9.5	0 mM	62.6 ± 0.0	
		40 mM	63.1 ± 0.1	
		100 mM	63.4 ± 0.1	
		500 mM	63.2 ± 0.0	
	11	0 mM	57.6 ± 0.3	
		40 mM	58.6 ± 0.1	
		100 mM	59.6 ± 0.0	
		500 mM	60.6 ± 0.7	
pET 151/D-TOPO DHDPS Q90L	4.5	0 mM	57.1 ± 0.1	
		40 mM	59.1 ± 0.1	
		100 mM	59.4 ± 0.1	
		500 mM	59.6 ± 0.0	
	7	0 mM	61.7 ± 0.0	70.8 ± 0.1
		40 mM	61.7 ± 0.0	70.2 ± 0.5
		100 mM	61.9 ± 0.1	69.7 ± 0.5
		500 mM	61.8 ± 0.1	
	9.5	0 mM	61.9 ± 0.1	74.3 ± 0.6
		40 mM	61.9 ± 0.1	74.7 ± 0.8
		100 mM	62.0 ± 0.0	74.6 ± 0.8
		500 mM	62.5 ± 0.2	
	11	0 mM	60.1 ± 0.1	
		40 mM	60.1 ± 0.1	
		100 mM	60.0 ± 0.1	
		500 mM	60.1 ± 0.1	
Cleaved pET 151/D-TOPO DHDPS Q90L	4.5	0 mM	59.2 ± 0.1	
		40 mM	60.4 ± 0.4	
		100 mM	60.5 ± 0.5	
		500 mM	60.4 ± 0.4	
	7	0 mM	60.1 ± 0.4	68.9 ± 0.0
		40 mM	61.2 ± 0.1	68.6 ± 0.0
		100 mM	61.4 ± 0.0	
		500 mM	61.6 ± 0.1	
	9.5	0 mM	61.1 ± 0.0	
		40 mM	61.1 ± 0.0	
		100 mM	61.3 ± 0.1	77.6 ± 0.2
		500 mM	61.4 ± 0.0	
	11	0 mM	56.6 ± 0.0	
		40 mM	56.7 ± 0.1	
		100 mM	56.9 ± 0.0	
		500 mM	57.5 ± 0.5	

Table A3.4 – The thermal denaturation temperatures (T_m) of DHDPS Q90L, pET M11 DHDPS Q90L, pET 151/D-TOPO DHDPS Q90L, and cleaved pET 151/D-TOPO DHDPS Q90L in 100 mM phosphate buffer at various pH and salt concentrations as monitored by SYPRO Orange fluorescence and as assessed from the derivation of the fluorescence data. The errors were calculated by identifying the maximal peak of the derivatives of each of the three replicates. The temperature at which each maximum occurred was established and the mean and standard error of the mean calculated.

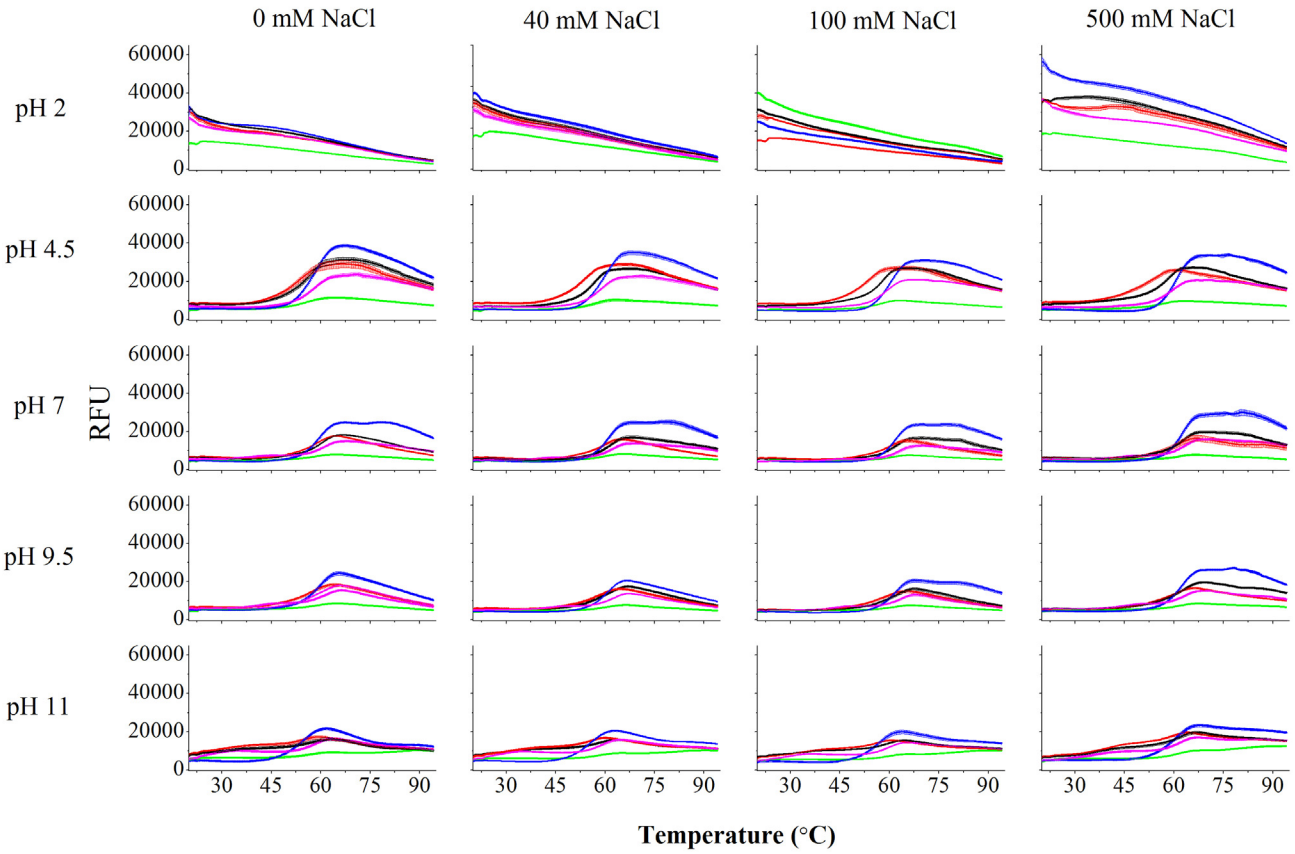


Figure A3.2 – The thermal denaturation of wild-type DHDPS (black), DHDPS A207V (red) pET M11 DHDPS A207V (green), pET 151/D-TOPO DHDPS A207V (blue) and cleaved pET 151/D-TOPO DHDPS A207V (pink) in 100 mM phosphate buffer at pH 2, pH 4.5, pH 7, pH 9.5 and pH 11 containing 0 mM NaCl, 40 mM NaCl, 100 mM NaCl and 500 mM NaCl as monitored by SYPRO Orange fluorescence. The data plotted are the mean increase in fluorescence, as monitored by the BioRad IQ5 ± SEM.

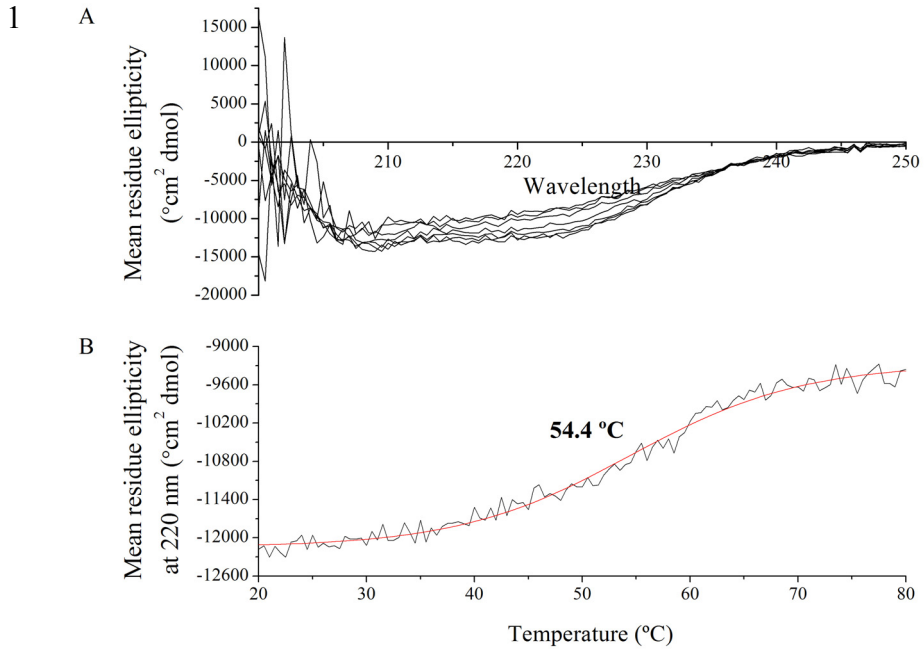
Variant	pH	Salt concentration	Mean thermal denaturation temperature 1 (°C) ± SEM	Mean thermal denaturation temperature 2 (°C) ± SEM
DHDPS A207V	4.5	0 mM	56.4 ± 0.1	
		40 mM	56.3 ± 0.0	
		100 mM	56.3 ± 0.0	
		500 mM	56.3 ± 0.0	
	7	0 mM	59.0 ± 0.3	
		40 mM	60.0 ± 0.2	
		100 mM	60.2 ± 0.3	
		500 mM	60.4 ± 0.2	
	9.5	0 mM	56.3 ± 0.0	
		40 mM	56.3 ± 0.0	
		100 mM	56.3 ± 0.0	

		500 mM	56.3 ± 0.0	
	11	0 mM	56.3 ± 0.0	
		40 mM	56.3 ± 0.0	
		100 mM	56.3 ± 0.0	
		500 mM	56.3 ± 0.0	
pET M11 DHDPS A207V	4.5	0 mM	56.1 ± 0.7	
		40 mM	57.6 ± 0.4	
		100 mM	57.6 ± 0.4	
		500 mM	58.1 ± 0.3	
	7	0 mM	57.9 ± 0.5	
		40 mM	58.2 ± 0.1	
		100 mM	59.8 ± 0.1	
		500 mM	60.4 ± 0.4	
	9.5	0 mM	59.6 ± 0.2	
		40 mM	59.9 ± 0.3	
		100 mM	60.5 ± 0.0	
		500 mM	60.5 ± 0.0	
	11	0 mM	54.5 ± 0.0	
		40 mM	54.5 ± 0.0	
		100 mM	54.5 ± 0.0	
		500 mM	54.5 ± 0.0	
pET 151/D-TOPO DHDPS A207V	4.5	0 mM	59.0 ± 0.0	
		40 mM	59.0 ± 0.0	
		100 mM	59.4 ± 0.4	
		500 mM	59.4 ± 0.4	
	7	0 mM	59.6 ± 0.2	74.5 ± 0.5
		40 mM	60.1 ± 0.1	74.3 ± 0.6
		100 mM	60.8 ± 0.5	74.7 ± 0.8
		500 mM	61.4 ± 0.0	74.6 ± 0.8
	9.5	0 mM	59.8 ± 0.1	
		40 mM	60.1 ± 0.1	
		100 mM	61.1 ± 0.2	76.6 ± 0.2
		500 mM	61.2 ± 0.1	72.8 ± 0.0
	11	0 mM	54.5 ± 0.0	
		40 mM	54.8 ± 0.0	
		100 mM	55.0 ± 0.2	
		500 mM	56.7 ± 0.1	
Cleaved pET 151/D-TOPO DHDPS A207V	4.5	0 mM	58.2 ± 0.1	
		40 mM	58.4 ± 0.0	
		100 mM	58.7 ± 0.0	
		500 mM	59.1 ± 0.2	
	7	0 mM	60.9 ± 0.1	
		40 mM	61.7 ± 0.3	
		100 mM	62.1 ± 0.1	
		500 mM	62.6 ± 0.2	
	9.5	0 mM	61.8 ± 0.4	
		40 mM	62.3 ± 0.0	
		100 mM	62.7 ± 0.1	
		500 mM	62.9 ± 0.0	

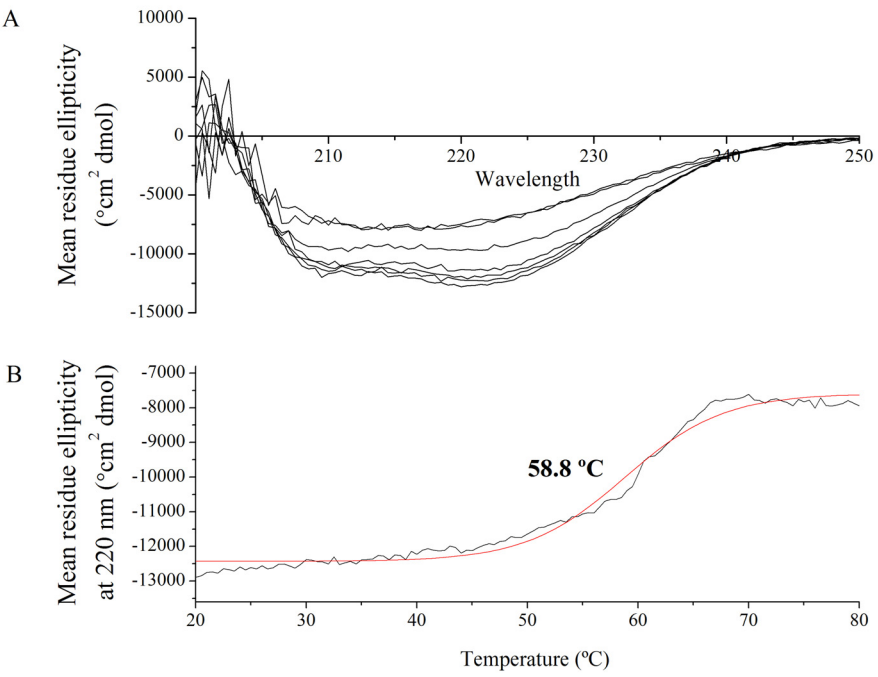
	11	0 mM	58.3 ± 0.1	
		40 mM	58.6 ± 0.1	
		100 mM	59.0 ± 0.3	
		500 mM	59.5 ± 0.4	

Table A3.5 – The thermal denaturation temperatures (T_m) of DHDPS A207V, pET M11 DHDPS A207V, pET 151/D-TOPO DHDPS A207V, and cleaved pET 151/D-TOPO DHDPS A207V in 100 mM phosphate buffer at various pH and salt concentrations as monitored by SYPRO Orange fluorescence and as assessed from the derivation of the fluorescence data. The errors were calculated by identifying the maximal peak of the derivatives of each of the three replicates. The temperature at which each maximum occurred was established and the mean and standard error of the mean calculated.

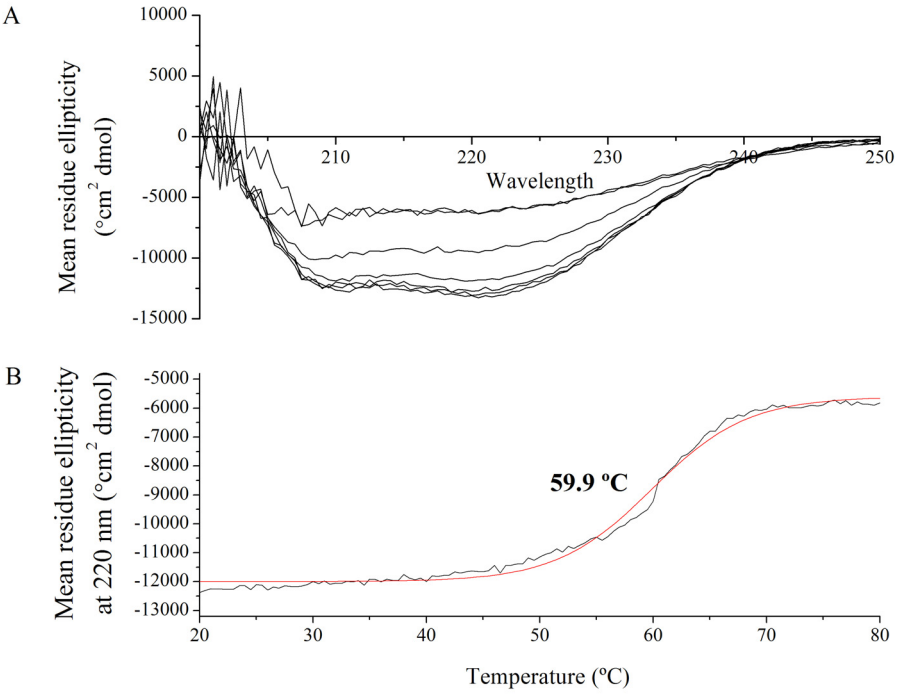
A3.3 CIRCULAR DICHROISM SPECTROSCOPY



2


A
Mean residue ellipticity (°cm^2 dmol)
Wavelength
B
Mean residue ellipticity at 220 nm (°cm^2 dmol)
58.8 ºC
Temperature (ºC)

3


A
Mean residue ellipticity (°cm^2 dmol)
Wavelength
B
Mean residue ellipticity at 220 nm (°cm^2 dmol)
59.9 ºC
Temperature (ºC)

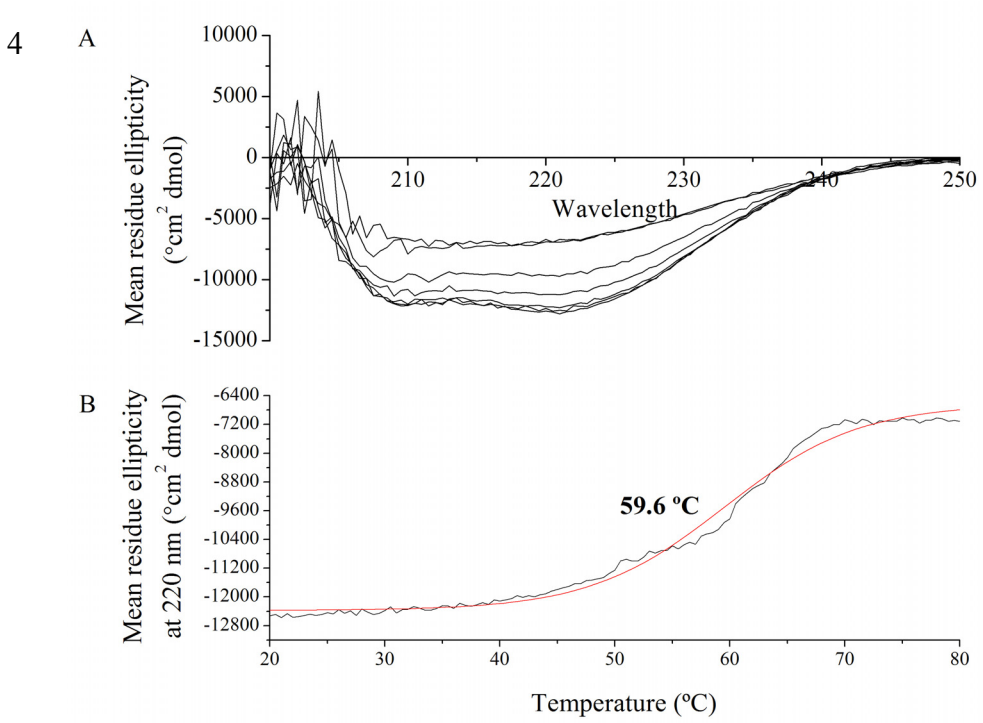
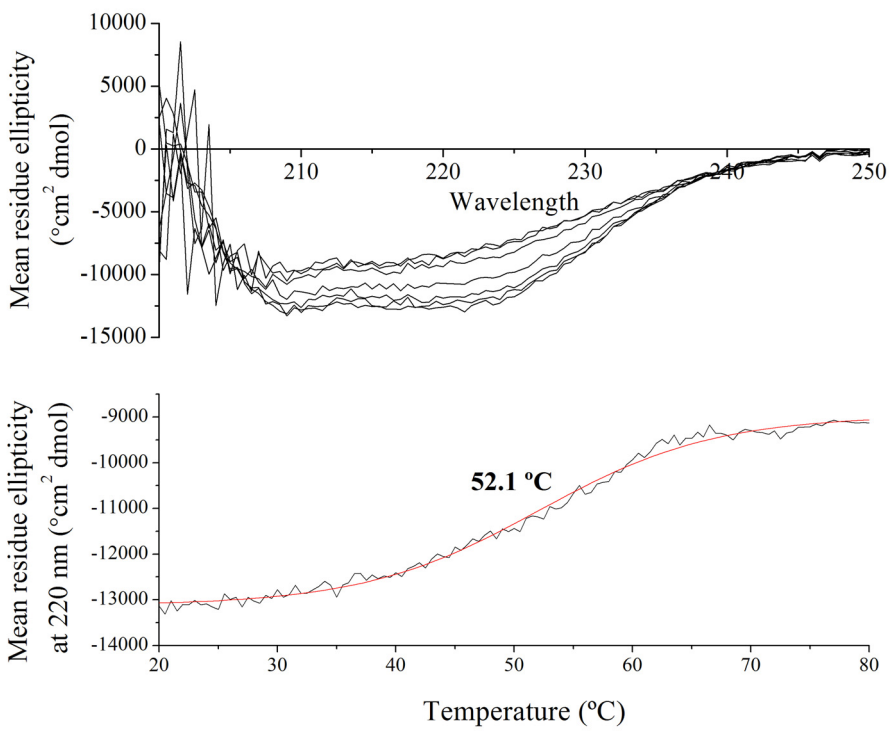
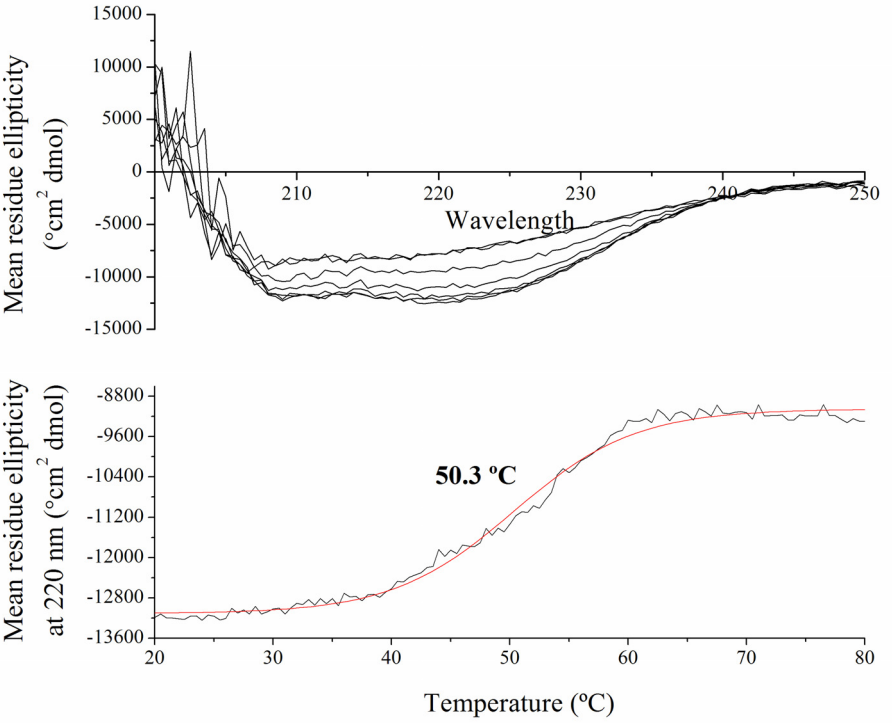


Figure A3.3 –Circular dichroism spectroscopy of **(1)** DHDPS Q90L **(2)** pET M11 DHDPS Q90L, **(3)** pET 151/D-TOPO DHDPS Q90L and **(4)** cleaved pET 151/D-TOPO DHDPS Q90L. **A)** Wavelength scans plotting mean residue ellipticity vs. wavelength at 20ºC, 30ºC, 40ºC, 50ºC, 60ºC, 70ºC and 80ºC. **B)** Mean residue ellipticity [θ] vs. temperature indicating melting of the protein. The red line corresponds to a sigmoid (Boltzmann) fit and the temperature shown is the midpoint of that fit.

1


A
Mean residue ellipticity (°cm^2 dmol)
10000
5000
0
-5000
-10000
-15000
210
220
230
240
250
Wavelength
B
Mean residue ellipticity at 220 nm (°cm^2 dmol)
-9000
-10000
-11000
-12000
-13000
-14000
52.1 °C
20
30
40
50
60
70
80
Temperature (°C)

2


A
Mean residue ellipticity (°cm^2 dmol)
15000
10000
5000
0
-5000
-10000
-15000
210
220
230
240
250
Wavelength
B
Mean residue ellipticity at 220 nm (°cm^2 dmol)
-8800
-9600
-10400
-11200
-12000
-12800
-13600
50.3 °C
20
30
40
50
60
70
80
Temperature (°C)

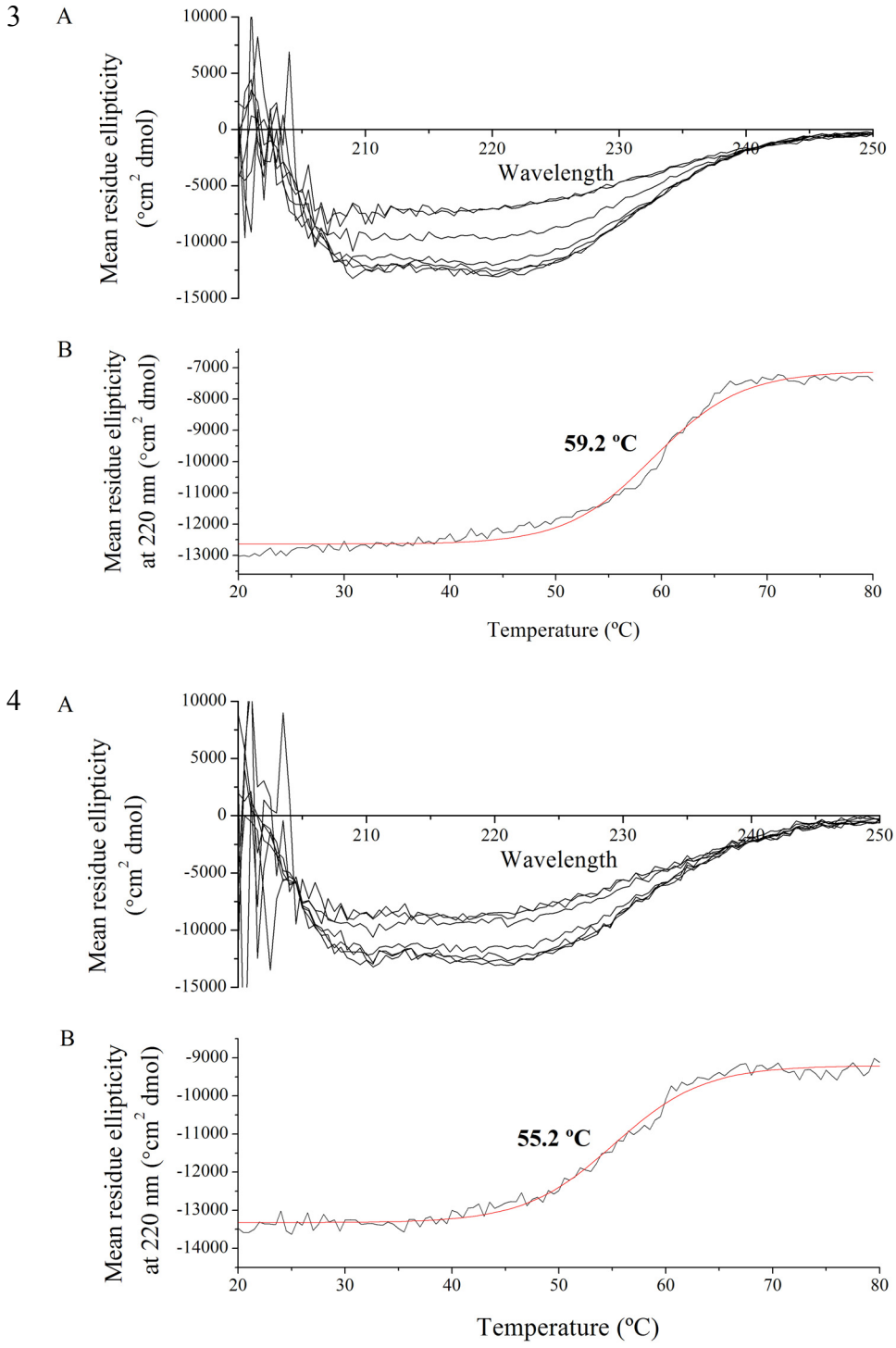


Figure A3.4 –Circular dichroism spectroscopy of **(1)** DHDPS A207V **(2)** pET 151/D-TOPO DHDPS A207V and **(3)** cleaved pET 151/D-TOPO DHDPS A207V. **A)** Wavelength scans plotting mean residue ellipticity vs. wavelength at 20°C, 30°C, 40°C, 50°C, 60°C, 70°C and 80°C. **B)** Mean residue ellipticity [θ] vs. temperature indicating melting of the protein. The red line corresponds to a sigmoid (Boltzmann) fit and the temperature shown is the midpoint of that fit.

A3.4 AMORPHOUS AGGREGATION

Variant	pH	Salt concentration (mM)	$Agg_{½}$ (minutes) ± SEM	Maximum absorbance (AU) ± SEM
DHDPS Q90L	2	0		0.15 ± 0.02
		20		0.12 ± 0.03
		40		0.15 ± 0.02
		100		0.21 ± 0.03
		500		0.95 ± 0.08
	4.5	0	9.76 ± 0.89	1.51 ± 0.03
		20	10.01 ± 0.47	1.58 ± 0.01
		40	10.93 ± 1.14	1.54 ± 0.06
		100	9.83 ± 0.34	1.56 ± 0.02
		500	9.43 ± 0.68	1.13 ± 0.05
	7	0	50.07 ± 1.76	0.93 ± 0.01
		20	52.45 ± 1.04	1.01 ± 0.02
		40	51.17 ± 0.93	1.04 ± 0.03
		100	43.40 ± 0.46	1.14 ± 0.01
		500	31.70 ± 0.54	1.08 ± 0.06
	9.5	0		0.10 ± 0.01
		20		0.09 ± 0.01
		40		0.28 ± 0.02
		100		0.45 ± 0.01
		500	31.20 ± 0.63	1.01 ± 0.03
	11	0		0.12 ± 0.01
		20		0.11 ± 0.01
		40		0.15 ± 0.01
		100		0.10 ± 0.02
		500		0.14 ± 0.04
pET M11 DHDPS Q90L	2	0		0.17 ± 0.07
		20		0.13 ± 0.08
		40		0.12 ± 0.01
		100		0.18 ± 0.02
		500	8.78 ± 1.58	1.44 ± 0.14
	4.5	0		1.62 ± 0.05
		20		1.62 ± 0.05
		40		1.60 ± 0.08
		100		1.56 ± 0.11
		500	34.63 ± 0.72	1.46 ± 0.04
	7	0	35.16 ± 0.77	1.19 ± 0.07
		20	39.24 ± 0.50	1.18 ± 0.08
		40	39.38 ± 0.50	0.99 ± 0.03
		100	40.00 ± 0.26	1.03 ± 0.06
		500	36.14 ± 0.28	1.02 ± 0.06
	9.5	0	48.86 ± 1.24	0.80 ± 0.10
		20	50.50 ± 1.54	0.88 ± 0.10
		40	45.26 ± 1.34	0.81 ± 0.09
		100	42.03 ± 0.78	0.74 ± 0.06

		500	34.49 ± 0.88	0.85 ± 0.06
	11	0		0.12 ± 0.02
		20		0.09 ± 0.00
		40		0.14 ± 0.01
		100		0.12 ± 0.01
		500		0.81 ± 0.11
pET 151/D-TOPO DHDPS Q90L	2	0		0.06 ± 0.01
		20		0.04 ± 0.01
		40		0.08 ± 0.01
		100		0.08 ± 0.02
		500	12.29 ± 0.71	0.88 ± 0.05
	4.5	0		1.31 ± 0.04
		20		1.18 ± 0.06
		40		1.24 ± 0.03
		100		1.22 ± 0.04
		500	32.76 ± 0.60	1.42 ± 0.14
	7	0	30.61 ± 0.84	0.72 ± 0.02
		20	32.98 ± 0.25	0.61 ± 0.03
		40	36.88 ± 0.30	0.65 ± 0.02
		100	36.38 ± 0.33	0.62 ± 0.02
		500	34.52 ± 0.51	0.61 ± 0.01
	9.5	0	36.69 ± 1.02	0.33 ± 0.01
		20	37.78 ± 0.25	0.27 ± 0.01
		40	39.67 ± 0.11	0.39 ± 0.01
		100	35.14 ± 0.28	0.32 ± 0.01
		500	32.23 ± 0.57	0.49 ± 0.02
	11	0		0.07 ± 0.01
		20		0.05 ± 0.00
		40		0.06 ± 0.01
		100		0.04 ± 0.01
		500		0.32 ± 0.01
Cleaved pET 151/D-TOPO DHDPS Q90L	2	0		0.07 ± 0.01
		20		0.11 ± 0.01
		40		0.12 ± 0.01
		100		0.16 ± 0.01
		500	12.23 ± 0.50	0.93 ± 0.02
	4.5	0		1.22 ± 0.04
		20		1.26 ± 0.03
		40		1.17 ± 0.05
		100	30.72 ± 0.14	1.16 ± 0.03
		500	32.78 ± 0.45	1.07 ± 0.02
	7	0	31.71 ± 0.69	0.78 ± 0.01
		20	35.92 ± 0.47	0.77 ± 0.02
		40	36.40 ± 0.33	0.71 ± 0.01
		100	37.66 ± 0.23	0.68 ± 0.02
		500	35.61 ± 0.29	0.73 ± 0.00
	9.5	0	40.83 ± 0.63	0.72 ± 0.05
		20	42.52 ± 0.16	0.58 ± 0.02
		40	40.47 ± 0.35	0.49 ± 0.02

		100	38.35 ± 0.18	0.50 ± 0.01
		500	33.13 ± 0.53	0.56 ± 0.01
	11	0		0.05 ± 0.01
		20		0.04 ± 0.01
		40		0.07 ± 0.01
		100		0.06 ± 0.01
		500		0.42 ± 0.01

Table A3.7 – The aggregation half life (*Agg½*) and maxima of DHDPS Q90L, pET M11 DHDPS Q90L, pET 151/D-TOPO DHDPS Q90L and cleaved pET 151/D-TOPO DHDPS Q90L as determined by the calculation of the midpoint of a sigmoid (Boltzmann) function fitted to the data in figure 4.19. Errors were calculated as per the CD data (chapter 2, figure 2.17). The maximum absorbance was the highest reading over the initial 90 minutes of the assay.

Variant	pH	Salt concentration (mM)	*Agg½* (minutes) ± SEM	Maximum absorbance (AU) ± SEM
DHDPS A207V	2	0		0.12 ± 0.01
		20		0.10 ± 0.02
		40		0.13 ± 0.02
		100		0.18 ± 0.03
		500	21.35 ± 0.26	0.44 ± 0.02
	4.5	0	15.74 ± 1.00	0.44 ± 0.01
		20	18.58 ± 0.71	0.42 ± 0.01
		40	17.80 ± 1.59	0.49 ± 0.01
		100	16.95 ± 0.97	0.48 ± 0.02
		500	15.25 ± 0.90	0.45 ± 0.03
	7	0	81.32 ± 1.06	0.24 ± 0.01
		20	53.68 ± 0.95	0.23 ± 0.01
		40	52.39 ± 0.83	0.23 ± 0.01
		100	48.89 ± 0.47	0.26 ± 0.01
		500	46.44 ± 0.54	0.26 ± 0.01
	9.5	0		0.09 ± 0.02
		20		0.09 ± 0.02
		40		0.12 ± 0.01
		100		0.10 ± 0.02
		500	40.16 ± 2.48	0.27 ± 0.02
	11	0		0.06 ± 0.02
		20		0.06 ± 0.02
		40		0.09 ± 0.01
		100		0.04 ± 0.01
		500		0.06 ± 0.02
pET M11 DHDPS A207V	2	0		0.12 ± 0.01
		20		0.08 ± 0.01
		40		0.13 ± 0.02
		100		0.18 ± 0.02
		500	22.05 ± 3.33	0.41 ± 0.04
	4.5	0	26.21 ± 0.23	0.44 ± 0.02

		20	26.81 ± 0.27	0.53 ± 0.03
		40	28.14 ± 0.16	0.54 ± 0.04
		100	26.70 ± 0.24	0.56 ± 0.03
		500	34.63 ± 0.67	0.52 ± 0.02
	7	0	38.59 ± 0.57	0.52 ± 0.01
		20	39.18 ± 0.82	0.50 ± 0.01
		40	41.04 ± 0.82	0.48 ± 0.02
		100	38.21 ± 0.66	0.48 ± 0.02
		500	52.21 ± 1.35	0.43 ± 0.01
	9.5	0	55.19 ± 1.84	0.23 ± 0.01
		20	49.34 ± 1.68	0.28 ± 0.03
		40	44.63 ± 0.69	0.36 ± 0.03
		100	37.58 ± 1.59	0.34 ± 0.01
		500		0.36 ± 0.02
	11	0		0.15 ± 0.01
		20		0.13 ± 0.00
		40		0.16 ± 0.01
		100		0.13 ± 0.01
		500		0.14 ± 0.01
pET 151/D-TOPO DHDPS A207V	2	0		0.02 ± 0.01
		20		0.00 ± 0.02
		40		0.06 ± 0.02
		100		0.09 ± 0.02
		500	15.49 ± 1.29	0.66 ± 0.05
	4.5	0	25.74 ± 0.83	0.96 ± 0.05
		20	28.02 ± 0.28	0.88 ± 0.05
		40	30.06 ± 0.52	0.93 ± 0.05
		100	29.08 ± 0.56	0.83 ± 0.03
		500	32.04 ± 0.99	0.83 ± 0.02
	7	0	33.97 ± 0.73	0.55 ± 0.02
		20	37.47 ± 0.60	0.52 ± 0.02
		40	37.97 ± 0.78	0.48 ± 0.03
		100	39.09 ± 0.34	0.53 ± 0.02
		500	38.41 ± 0.74	0.50 ± 0.01
	9.5	0	37.67 ± 0.67	0.45 ± 0.02
		20	40.68 ± 0.29	0.41 ± 0.01
		40	41.64 ± 0.51	0.47 ± 0.03
		100	34.90 ± 0.34	0.43 ± 0.01
		500	35.68 ± 0.37	0.38 ± 0.01
	11	0		0.02 ± 0.00
		20		0.02 ± 0.01
		40		0.08 ± 0.01
		100		0.03 ± 0.02
		500		0.38 ± 0.02

Table A3.8 – The aggregation half life ($Agg_{½}$) and maxima of DHDPS A207V, pET M11 DHDPS A207V, pET 151/D-TOPO DHDPS A207V and cleaved pET 151/D-TOPO DHDPS A207V as determined by the calculation of the midpoint of a sigmoid (Boltzmann) function fitted to the data in figure 4.21. Errors were

calculated as per the CD data (chapter 2, figure 2.17). The maximum absorbance was the highest reading over the initial 90 minutes of the assay.

	Degrees of freedom	Sum of squares	Mean squares	F value	Pr (>F)
Protein	7	7423	1060	399.32	$< 2.2 \times 10^{-16}$
Salt	4	1027	257	96.68	$< 2.2 \times 10^{-16}$
Protein:salt	28	2403	86	32.32	$< 2.2 \times 10^{-16}$
Residuals	200	531.1	2.7		

Table A3.9 - Analysis of deviance table for the statistical analysis of the $Agg_{½}$ of DHDPS Q90L, A207V and their polyhistidine tagged variants.

	Degrees of freedom	Sum of squares	Mean squares	F value	Pr (>F)
Protein	7	50.8	7.30	383.7	$< 2.2 \times 10^{-16}$
Salt	4	8.4	2.10	111	$< 2.2 \times 10^{-16}$
pH	4	120.7	30.19	1596.4	$< 2.2 \times 10^{-16}$
Protein:salt	28	3.7	0.13	6.9	$< 2.2 \times 10^{-16}$
Protein:pH	28	25.1	0.90	47.5	$< 2.2 \times 10^{-16}$
salt:pH	16	17.9	1.12	59.2	$< 2.2 \times 10^{-16}$
Protein:pH:salt	112	9.6	0.09	4.5	$< 2.2 \times 10^{-16}$
Residuals	1000	18.9	0.02		

Table A3.10 - Analysis of deviance table for the statistical analysis of the maximum absorbance of DHDPS Q90L, A207V and their polyhistidine tagged variants.

A3.5 β-SHEET-SPECIFIC AGGREGATION

Variant	pH	Salt concentration (mM)	β-$agg_{½}$ (minutes) ± SEM	Maximum fluorescence (RFU) ± SEM
DHDPS Q90L	2	0		-120 ± 24
		40		103 ± 27
		500		127 ± 16
	4.5	0		736 ± 59
		40		754 ± 73
		500		1400 ± 78
	7	0	14.53 ± 0.89	896 ± 39
		40	76.64 ± 6.55	1040 ± 47
		500	92.95 ± 2.51	1299 ± 91
	9.5	0	13.34 ± 1.26	691 ± 17

		40	17.46 ± 5.66	576 ± 35
		500	72.79 ± 1.80	670 ± 29
	11	0		252 ± 25
		40		233 ± 30
		500		539 ± 38
pET M11 DHDPS Q90L	2	0		41 ± 9
		40		63 ± 11
		500		986 ± 139
	4.5	0		6428 ± 582
		40		5709 ± 427
		500		6045 ± 374
	7	0	78.38 ± 0.53	13140 ± 1370
		40	84.31 ± 0.70	11958 ± 820
		500	74.59 ± 4.26	9192 ± 816
	9.5	0	45.53 ± 1.86	1013 ± 126
		40	45.44 ± 1.63	1196 ± 141
		500	40.38 ± 5.78	2537 ± 365
	11	0		6 ± 21
		40		19 ± 31
		500		57 ± 66
pET 151/D-TOPO DHDPS Q90L	2	0		185 ± 134
		40		116 ± 22
		500		763 ± 19
	4.5	0		4552 ± 203
		40		4943 ± 174
		500		6103 ± 430
	7	0	69.88 ± 1.89	9702 ± 180
		40	76.96 ± 0.56	8278 ± 326
		500	76.57 ± 1.78	5952 ± 289
	9.5	0	47.70 ± 0.60	1478 ± 116
		40	51.20 ± 0.37	1645 ± 58
		500	52.52 ± 0.57	2526 ± 75
	11	0		41 ± 21
		40		10 ± 13
		500		92 ± 8
Cleaved pET 151/D-TOPO DHDPS Q90L	2	0		-172 ± 12
		40		-145 ± 6
		500		373 ± 66
	4.5	0		2754 ± 98
		40		2613 ± 211
		500		4453 ± 141
	7	0	72.24 ± 0.55	7545 ± 375
		40	73.83 ± 1.33	5761 ± 296
		500	76.39 ± 1.33	5302 ± 232
	9.5	0	54.79 ± 0.47	2354 ± 55
		40	54.21 ± 0.30	1909 ± 47
		500	49.76 ± 1.20	1937 ± 89

	11	0		-72 ± 9
		40		-52 ± 15
		500		-65 ± 6

Table A3.11 – The speed of β-sheet specific aggregation (β-$agg_{½}$) of DHDPS Q90L, pET M11 DHDPS Q90L, pET 151/D-TOPO DHDPS Q90L and cleaved pET 151/D-TOPO DHDPS Q90L as determined by the calculation of the $t_{½}$ of a sigmoid (Boltzmann) function fitted to the ThT fluorescence data (chapter 4, figure 4.23). Errors were calculated as per the CD data (chapter 2, figure 2.17). The maximum absorbance was the highest reading over the initial 300 minutes of the assay.

Variant	pH	Salt concentration (mM)	β-$agg_{½}$ (minutes) ± SEM	Maximum fluorescence (RFU) ± SEM
DHDPS A207V	2	0		-44 ± 68
		40		-73 ± 13
		500		272 ± 69
	4.5	0		903 ± 88
		40		956 ± 134
		500		1775 ± 178
	7	0	37.58 ± 2.31	1407 ± 52
		40	69.93 ± 2.36	1619 ± 46
		500	69.33 ± 2.78	1755 ± 158
	9.5	0	32.40 ± 0.27	723 ± 36
		40	32.77 ± 0.44	754 ± 31
		500	57.68 ± 1.83	1024 ± 64
	11	0		126 ± 28
		40		254 ± 46
		500		276 ± 35
pET M11 DHDPS A207V	2	0		-196 ± 15
		40		-222 ± 17
		500		-79 ± 108
	4.5	0		253 ± 36
		40		359 ± 192
		500		542 ± 57
	7	0	53.03 ± 5.12	698 ± 33
		40	68.74 ± 2.09	641 ± 101
		500	68.20 ± 4.05	491 ± 36
	9.5	0	47.20 ± 1.41	341 ± 76
		40	50.94 ± 1.08	311 ± 105
		500	26.02 ± 1.89	799 ± 447
	11	0		-127 ± 28
		40		-78 ± 25
		500		-76 ± 37
pET 151/D-TOPO DHDPS A207V	2	0		146 ± 13
		40		149 ± 21
		500		801 ± 69
	4.5	0		3741 ± 73
		40		3912 ± 119

		500		5808 ± 361
	7	0	72.06 ± 1.09	6981 ± 358
		40	79.25 ± 1.27	6222 ± 316
		500	75.62 ± 1.89	4836 ± 305
	9.5	0	54.75 ± 1.22	1955 ± 88
		40	59.74 ± 1.00	2052 ± 58
		500	57.30 ± 0.93	2168 ± 88
	11	0		1 ± 26
		40		82 ± 18
		500		186 ± 15
Cleaved pET 151/D-TOPO DHDPS A207V	2	0		-140 ± 39
		40		-265 ± 12
		500		-142 ± 16
	4.5	0		805 ± 27
		40		831 ± 24
		500		1097 ± 30
	7	0	66.66 ± 1.62	2459 ± 68
		40	69.90 ± 1.33	2311 ± 34
		500	73.67 ± 2.48	1881 ± 31
	9.5	0	65.00 ± 0.52	34 ± 14
		40	71.53 ± 0.26	-49 ± 17
		500	74.77 ± 0.53	342 ± 16
	11	0		-114 ± 7
		40		-73 ± 2
		500		53 ± 11

Table A3.12 – The speed of β-sheet specific aggregation (β-$agg_{½}$) of DHDPS A207V, pET M11 DHDPS A207V, pET 151/D-TOPO DHDPS A207V and cleaved pET 151/D-TOPO DHDPS A207V as determined by the calculation of the $t_{½}$ of a sigmoid (Boltzmann) function fitted to the ThT fluorescence data (chapter 4, figure 4.24). Errors were calculated as per the CD data (chapter 2, figure 2.17). The maximum absorbance was the highest reading over the initial 300 minutes of the assay.

	Degrees of freedom	**Sum of squares**	**Mean squares**	**F value**	**Pr (>F)**
Protein	7	12024	1718	42	$< 2.2 \times 10^{-16}$
Salt	2	11137	5568	135	$< 2.2 \times 10^{-16}$
Protein:salt	14	16763	1197	29	$< 2.2 \times 10^{-16}$
Residuals	120	4939	41		

Table A3.13 - Analysis of deviance table for the statistical analysis of the β-$agg_{½}$ of DHDPS Q90L, A207V and their polyhistidine tagged variants. Note that the β-$agg_{½}$ was only calculated for one pH due to incomplete data sets for other pH values. As discussed in chapter 2, section 2.4.8, salt was included in the model because it had an effect on the fluorescence however this was not relevant due to its affect on ThT binding. The main effect of protein was the independent effect after controlling between the differences across pH and salt concentrations.

	Degrees of freedom	Sum of squares	Mean squares	F value	Pr (>F)
Protein	8	1220183979	152522997	319	$< 2.2 \times 10^{-16}$
pH	4	2454993755	613748439	1282	$< 2.2 \times 10^{-16}$
Salt	2	509566	254783	0.53	5.8×10^{-1}
Protein:pH	32	1447220110	45225628	94	$< 2.2 \times 10^{-16}$
Protein:salt	16	39911476	2494467	5.2	2.1×10^{-10}
pH:salt	8	2184044	273005	0.57	8.0×10^{-1}
Protein:pH:salt	64	42263245	660363	1.4	3.1×10^{-2}
Residuals	675	323030385	478564		

Table A3.14 - Analysis of deviance table for the statistical analysis of the maximum fluorescence of DHDPS Q90L, A207V and their polyhistidine tagged variants. As discussed in chapter 2, section 2.4.8, pH and salt were included in the model because they had an effect on the fluorescence however this was not relevant due to their affect on ThT binding.

APPENDIX 4

SUPPLEMENTARY INFORMATION FOR CHAPTER 5

A4.1 ALGORITHMIC PREDICTION

	Wild-type DHDPS	DHDPS Y107W	pET M11 DHDPS Y107W	pET 151/D-TOPO DHDPS Y107W	Cleaved pET 151/D-TOPO DHDPS Y107W	DHDPS Y107F
Overall hydrophobicity at pH 7	95.38	94.72	105.17	98.37	94.45	94.58
Overall charge at pH 7	-5	-5	-7	-7	-6	-5
Overall β-sheet propensity	1090.06	1090.02	1186.24	1195.37	1105.65	1090.15
Overall α-helical propensity	1133.27	1133.14	1234.22	1250.15	1151.16	1133.33
Number of hydrophobic/hydrophilic patterns	2	2	2	2	2	2
Absolute amyloid aggregation rate	-4.07	-4.06	-4.25	-4.26	-4.14	-4.06
Intrinsic aggregation propensity	-6.12	-6.11	-6.73	-6.86	-6.28	-6.11
Average ± SD	-5.46 ± 1.45	-5.45 ± 1.45	'-5.74 ± 1.51	-5.81 ± 1.53	-5.51 ± 1.47	-5.46 ± 1.46
Zagg score	-0.458258	-0.45746	-0.654976	-0.68411	-0.523118	-0.449314

Table A4.1 – The results of the Zyggregator algorithm for wild-type DHDPS, DHDPS Y107W, pET M11 DHDPS Y107W, pET 155/D-TOPO DHDPS Y107W, cleaved pET 151/D-TOPO DHDPS Y107W, and DHDPS Y107F.

	Wild-type DHDPS	pET M11 DHDPS L197Y	pET 151/D-TOPO DHDPS L197Y	Cleaved pET 151/D-TOPO DHDPS L197Y
Overall hydrophobicity at pH 7	95.38	106.18	99.38	95.46
Overall charge at pH 7	-5	-7	-7	-6
Overall β-sheet propensity	1090.06	1186.31	1195.43	1105.72
Overall α-helical propensity	1133.27	1234.04	1249.97	1159.98
Number of hydrophobic/hydrophilic patterns	2	2	2	2
Absolute amyloid aggregation rate	-4.07	-4.25	-4.26	-4.14
Intrinsic aggregation propensity	-6.12	-6.73	-6.86	-6.28
Average ± SD	-5.46 ± 1.45	-5.74 ± 1.51	-5.81 ± 1.54	-5.52 ± 1.47
Zagg score	-0.458258	-0.658108	-0.68622	-0.519414

Table A4.2 – The results of the Zyggregator algorithm for wild-type DHDPS, pET M11 DHDPS L197Y, pET 155/D-TOPO DHDPS L197Y and cleaved pET 151/D-TOPO DHDPS L197Y.

A4.2 DIFFERENTIAL SCANNING FLUORIMETRY

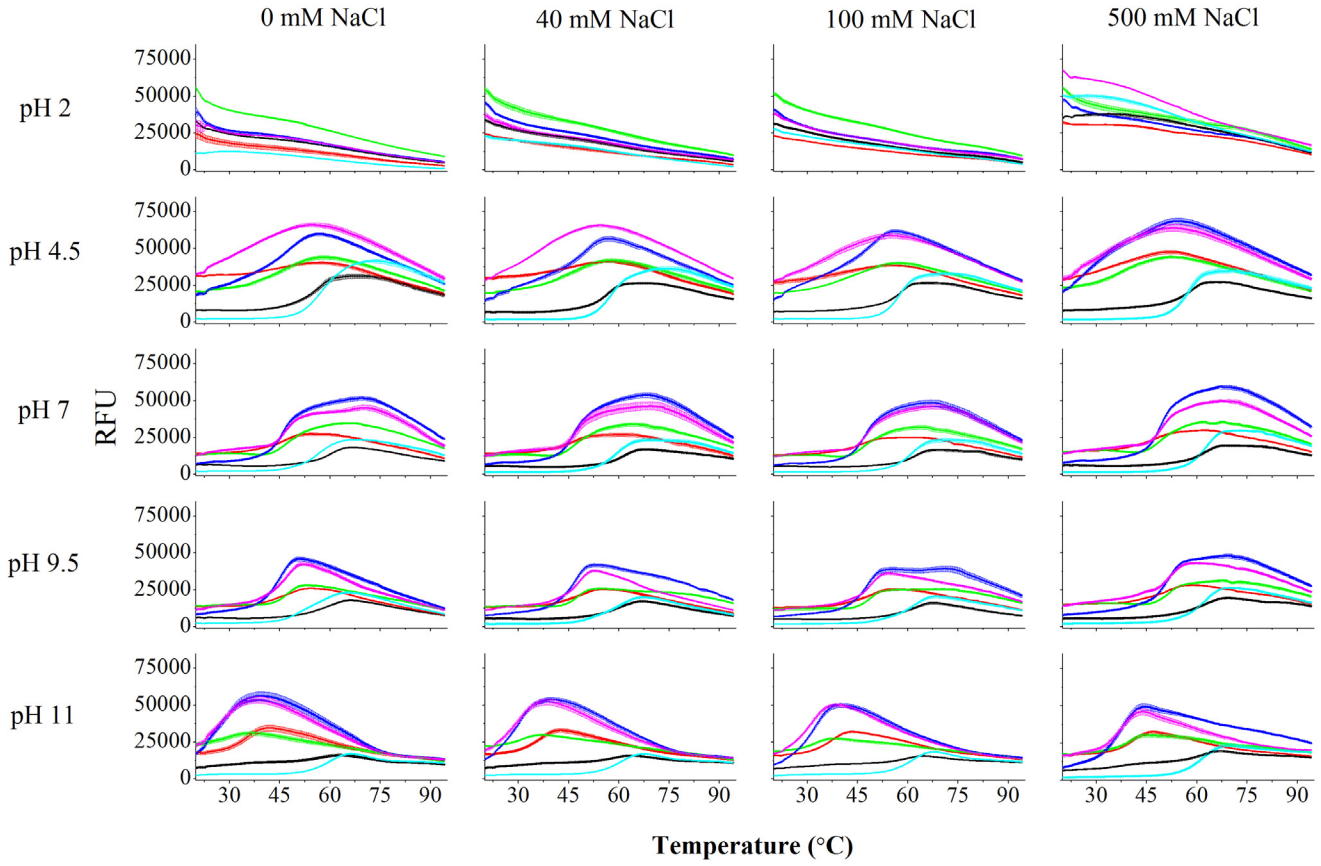


Figure A4.1 – The thermal denaturation of wild-type *E. coli* DHDPS (black), DHDPS Y107W (red) pET M11 DHDPS Y107W (green), pET 151/D-TOPO DHDPS Y107W (dark blue), cleaved pET 151/D-TOPO DHDPS Y107W (pink) and DHDPS Y107F (light blue) in 100 mM phosphate buffer at pH 2, pH 4.5, pH 7, pH 9.5 and pH 11 containing 0 mM NaCl, 40 mM NaCl, 100 mM NaCl and 500 mM NaCl as monitored by SYPRO Orange fluorescence. The data plotted are the increase in fluorescence, as monitored by the BioRad IQ5.

Variant	pH	Salt concentration	Mean thermal denaturation temperature 1 (°C) ± SEM	Mean thermal denaturation temperature 2 (°C) ± SEM
DHDPS Y107W	4.5	0 mM	45.5 ± 0.0	
		40 mM	46.5 ± 1.0	
		100 mM	46.5 ± 1.0	
		500 mM	42.3 ± 3.2	
	7	0 mM	47.9 ± 0.5	
		40 mM	47.3 ± 0.5	
		100 mM	47.0 ± 0.0	
		500 mM	47.5 ± 0.4	
	9.5	0 mM	47.5 ± 0.0	
		40 mM	49.0 ± 0.7	
		100 mM	49.7 ± 0.0	
		500 mM	50.6 ± 0.5	
	11	0 mM	34.5 ± 0.1	
		40 mM	36.4 ± 1.0	
		100 mM	36.4 ± 1.0	
		500 mM	37.4 ± 0.0	
pET M11 DHDPS Y107W	4.5	0 mM	37.9 ± 0.3	
		40 mM	42.1 ± 0.2	
		100 mM	42.3 ± 0.2	
		500 mM	42.1 ± 0.2	
	7	0 mM	46.7 ± 0.0	
		40 mM	46.7 ± 0.0	
		100 mM	46.7 ± 0.0	
		500 mM	47.6 ± 0.0	
	9.5	0 mM	47.6 ± 0.0	
		40 mM	47.7 ± 0.1	
		100 mM	47.9 ± 0.0	
		500 mM	49.5 ± 0.1	
	11	0 mM	29.9 ± 0.6	
		40 mM	30.5 ± 0.0	
		100 mM	30.5 ± 0.0	
		500 mM	40.2 ± 0.1	
pET 151/D-TOPO DHDPS Y107W	4.5	0 mM	47.1 ± 0.5	
		40 mM	49.4 ± 0.0	
		100 mM	47.8 ± 0.8	
		500 mM	26.9 ± 0.6	55.1 ± 0.0
	7	0 mM	45.5 ± 0.0	61.3 ± 0.8
		40 mM	46.1 ± 0.3	61.3 ± 0.7
		100 mM	45.9 ± 0.3	60.2 ± 0.0
		500 mM	47.0 ± 0.2	60.2 ± 0.0
	9.5	0 mM	43.4 ± 0.0	
		40 mM	46.2 ± 0.1	
		100 mM	47.0 ± 0.3	68.4 ± 0.1
		500 mM	47.1 ± 0.4	
	11	0 mM	29.4 ± 0.1	
		40 mM	30.4 ± 0.4	
		100 mM	30.8 ± 0.0	

		500 mM	30.8 ± 0.0	
Cleaved pET 151/D-TOPO DHDPS Y107W	4.5	0 mM	23.0 ± 0.0	
		40 mM	23.0 ± 0.0	
		100 mM	23.0 ± 0.0	
		500 mM	23.0 ± 0.0	
	7	0 mM	45.8 ± 0.2	66.8 ± 0.0
		40 mM	46.1 ± 0.2	55.7 ± 0.3
		100 mM	47.2 ± 0.6	61.0 ± 1.9
		500 mM	48.0 ± 0.1	59.0 ± 0.0
	9.5	0 mM	47.9 ± 0.0	
		40 mM	47.4 ± 0.3	
		100 mM	48.1 ± 0.2	
		500 mM	48.5 ± 0.3	
	11	0 mM	23.3 ± 0.0	
		40 mM	23.2 ± 0.1	
		100 mM	23.2 ± 0.1	
		500 mM	30.7 ± 0.1	
DHDPS Y107F	4.5	0 mM	59.2 ± 0.2	
		40 mM	58.6 ± 0.6	
		100 mM	57.4 ± 0.1	
		500 mM	57.6 ± 0.1	
	7	0 mM	56.3 ± 0.2	
		40 mM	58.6 ± 0.3	
		100 mM	59.5 ± 0.1	
		500 mM	60.7 ± 0.7	
	9.5	0 mM	51.1 ± 0.1	
		40 mM	59.2 ± 0.4	
		100 mM	59.6 ± 0.0	
		500 mM	62.3 ± 0.0	
	11	0 mM	59.7 ± 0.1	
		40 mM	59.8 ± 0.1	
		100 mM	60.3 ± 0.4	
		500 mM	61.7 ± 0.3	

Table A4.3 – The thermal denaturation temperatures (T_m) of DHDPS Y107W, pET M11 DHDPS Y107W, pET 151/D-TOPO DHDPS Y107W, cleaved pET 151/D-TOPO DHDPS Y107W and DHDPS Y107F in 100 mM phosphate buffer at various pH and salt concentrations as monitored by SYPRO Orange fluorescence and as assessed from the derivation of the fluorescence data. The errors were calculated by identifying the maximal peak of the derivatives of each of the three replicates. The temperature at which each maximum occurred was established and the mean and standard error of the mean calculated.

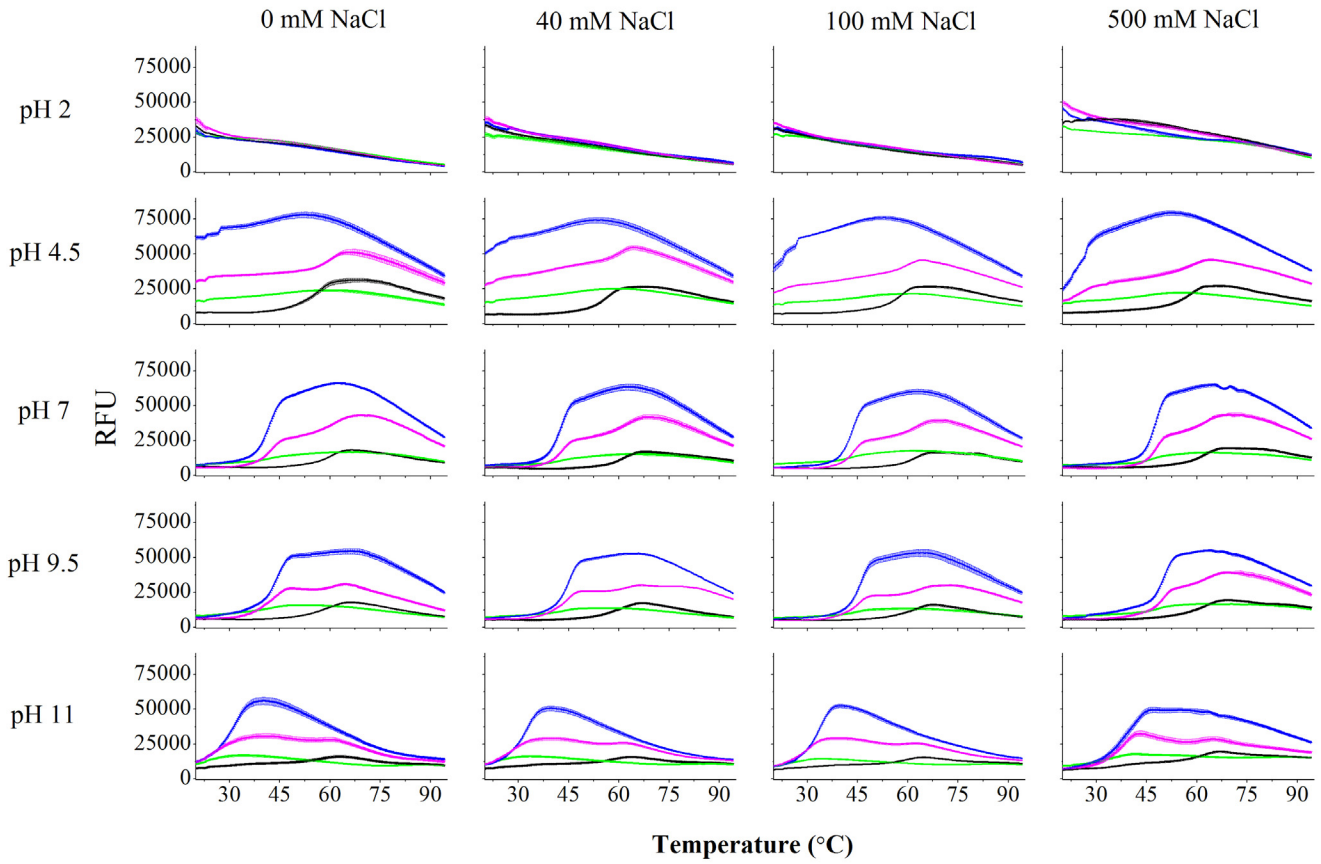


Figure A4.2 – The thermal denaturation of wild-type DHDPS (black), pET M11 DHDPS L197Y (green), pET 151/D-TOPO DHDPS L197Y (blue) and cleaved pET 151/D-TOPO DHDPS L197Y (pink) in 100 mM phosphate buffer at pH 2, pH 4.5, pH 7, pH 9.5 and pH 11 containing 0 mM NaCl, 40 mM NaCl, 100 mM NaCl and 500 mM NaCl as monitored by SYPRO Orange fluorescence. The data plotted are the increase in fluorescence, as monitored by the BioRad IQ5.

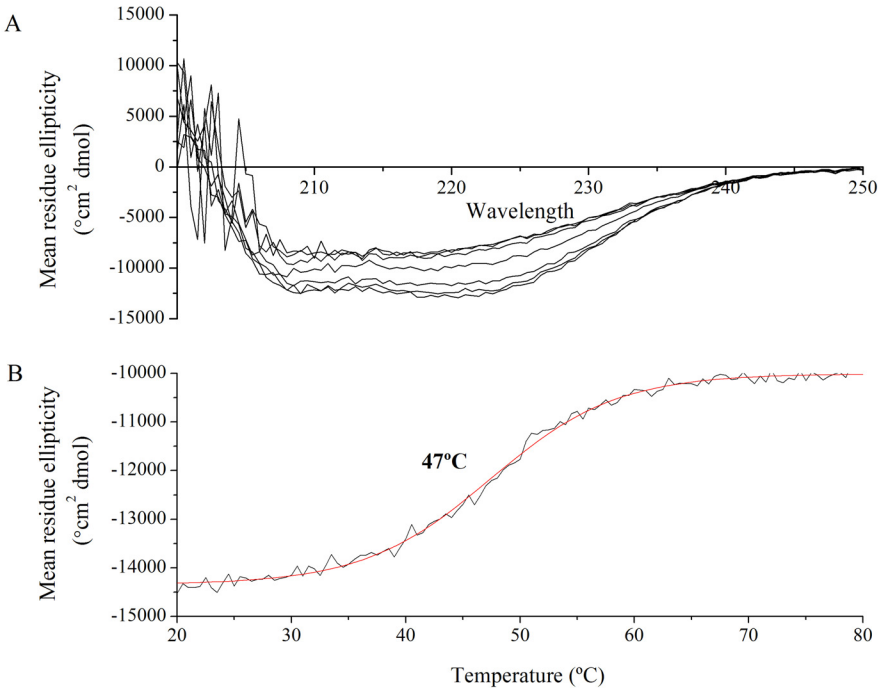
Variant	pH	Salt concentration	Mean thermal denaturation temperature 1 (°C) ± SEM	Mean thermal denaturation temperature 2 (°C) ± SEM
pET M11 DHDPS L197Y	4.5	0 mM	45.8 ± 0.3	
		40 mM	48.9 ± 0.2	
		100 mM	48.0 ± 0.1	
		500 mM	47.1 ± 0.4	
	7	0 mM	42.2 ± 0.3	
		40 mM	42.6 ± 0.2	
		100 mM	42.9 ± 0.1	
		500 mM	43.4 ± 0.2	
	9.5	0 mM	42.8 ± 0.2	

		40 mM	43.4 ± 0.5	
		100 mM	44.7 ± 0.2	
		500 mM	44.7 ± 0.1	
	11	0 mM	23.0 ± 0.0	
		40 mM	23.0 ± 0.0	
		100 mM	23.0 ± 0.0	
		500 mM	23.0 ± 0.0	
pET 151/D-TOPO DHDPS L197Y	4.5	0 mM	44.0 ± 0.0	
		40 mM	44.0 ± 0.0	
		100 mM	44.0 ± 0.0	
		500 mM	41.8 ± 2.2	
	7	0 mM	42.0 ± 0.1	57.1 ± 0.4
		40 mM	42.4 ± 0.4	56.7 ± 0.4
		100 mM	42.9 ± 0.4	56.7 ± 0.4
		500 mM	43.6 ± 0.1	56.3 ± 0.0
	9.5	0 mM	44.2 ± 0.2	58.9 ± 1.7
		40 mM	44.7 ± 0.1	57.7 ± 1.0
		100 mM	45.0 ± 0.2	58.0 ± 0.8
		500 mM	45.8 ± 0.3	58.9 ± 0.7
	11	0 mM	32.0 ± 0.0	
		40 mM	32.0 ± 0.0	
		100 mM	32.5 ± 0.4	
		500 mM	33.3 ± 0.1	
Cleaved pET 151/D-TOPO DHDPS L197Y	4.5	0 mM	61.0 ± 0.0	
		40 mM	61.1 ± 0.0	
		100 mM	61.1 ± 0.0	
		500 mM	60.2 ± 0.9	
	7	0 mM	42.8 ± 0.0	61.5 ± 0.1
		40 mM	43.2 ± 0.2	61.6 ± 0.1
		100 mM	43.4 ± 0.0	61.4 ± 0.2
		500 mM	43.7 ± 0.0	62.2 ± 0.8
	9.5	0 mM	43.7 ± 0.0	60.8 ± 0.5
		40 mM	44.4 ± 0.1	61.3 ± 0.1
		100 mM	44.7 ± 0.2	61.2 ± 0.1
		500 mM	45.3 ± 0.4	61.3 ± 0.1
	11	0 mM	23.0 ± 0.0	57.2 ± 0.6
		40 mM	23.0 ± 0.0	58.6 ± 0.1
		100 mM	23.0 ± 0.0	58.5 ± 0.1
		500 mM	30.1 ± 0.9	59.3 ± 0.6

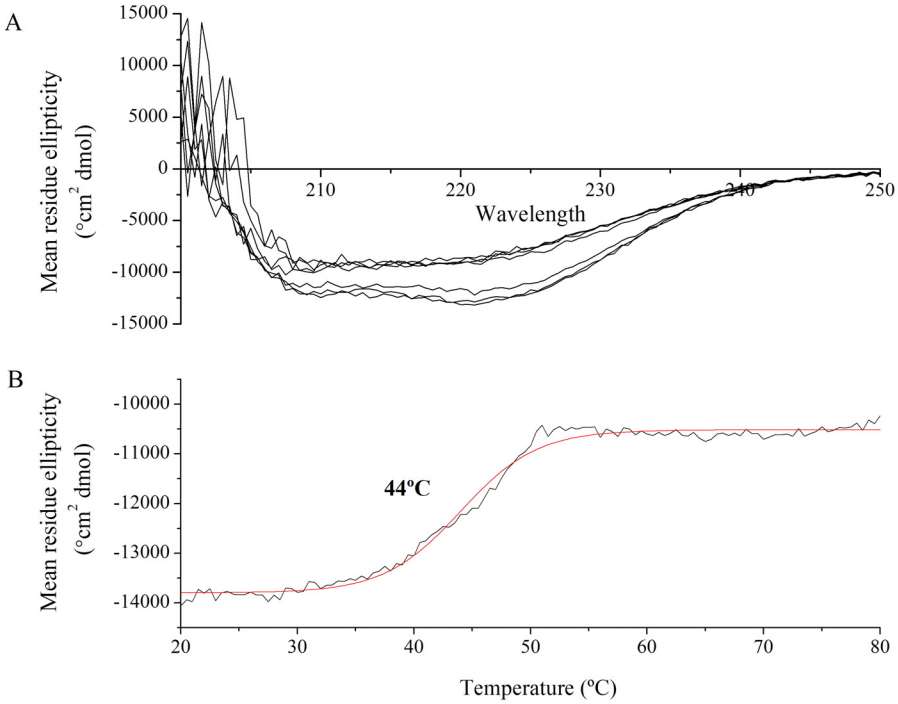
Table A4.4 – The thermal denaturation temperatures (T_m) of pET M11 DHDPS L197Y, pET 151/D-TOPO DHDPS L197Y, cleaved pET 151/D-TOPO and DHDPS L197Y Y107F in 100 mM phosphate buffer at various pH and salt concentrations as monitored by SYPRO Orange fluorescence and as assessed from the derivation of the fluorescence data. The errors were calculated by identifying the maximal peak of the derivatives of each of the three replicates. The temperature at which each maximum occurred was established and the mean and standard error of the mean calculated.

A4.3 CIRCULAR DICHROISM SPECTROSCOPY

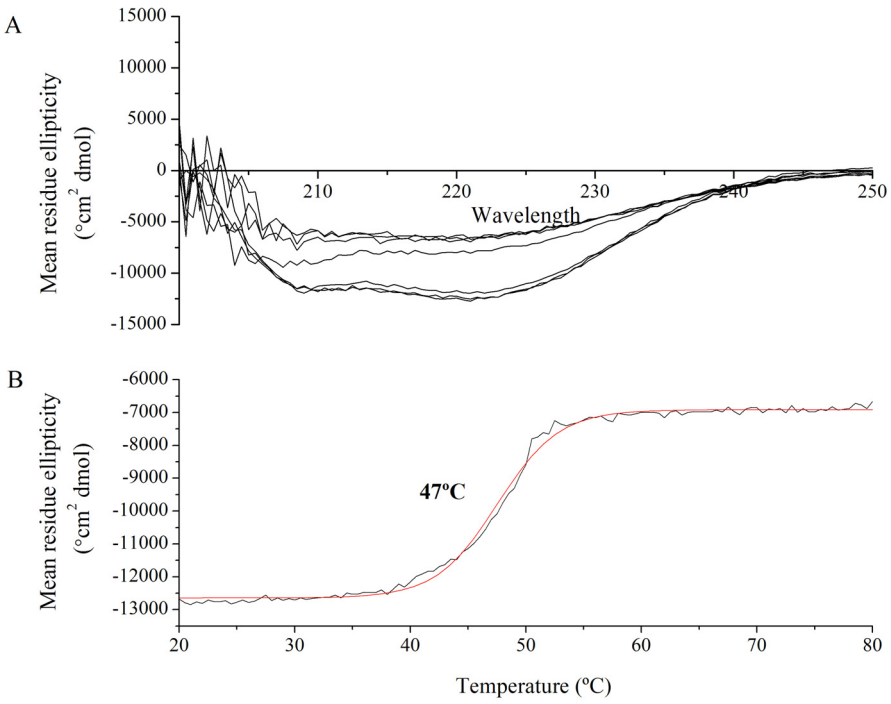
1



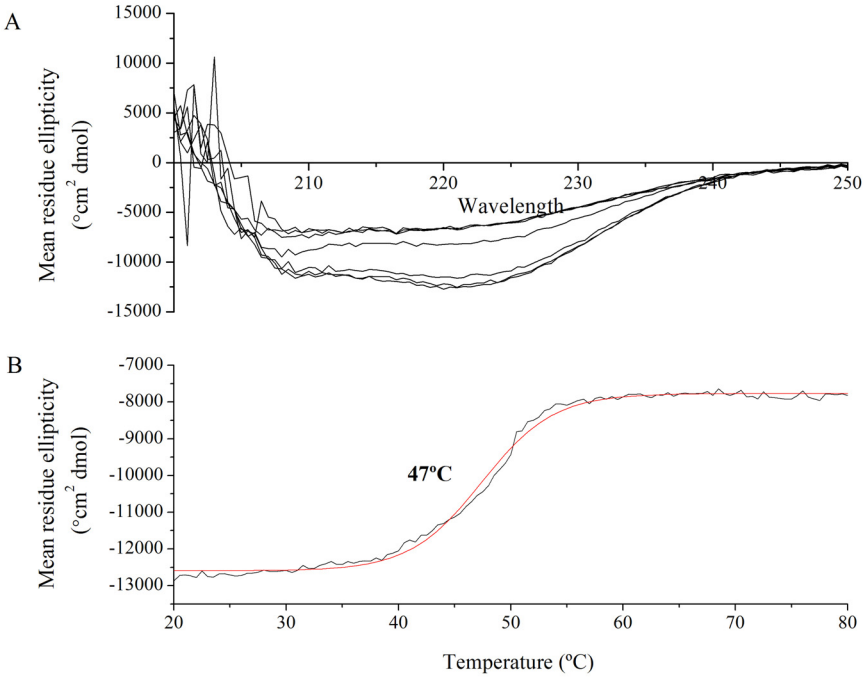
2



3


A
Mean residue ellipticity
(°cm^2 dmol)
15000
10000
5000
0
-5000
-10000
-15000
210
220
230
240
250
Wavelength
B
Mean residue ellipticity
(°cm^2 dmol)
-6000
-7000
-8000
-9000
-10000
-11000
-12000
-13000
47°C
20
30
40
50
60
70
80
Temperature (°C)

4


A
Mean residue ellipticity
(°cm^2 dmol)
15000
10000
5000
0
-5000
-10000
-15000
210
220
230
240
250
Wavelength
B
Mean residue ellipticity
(°cm^2 dmol)
-7000
-8000
-9000
-10000
-11000
-12000
-13000
47°C
20
30
40
50
60
70
80
Temperature (°C)

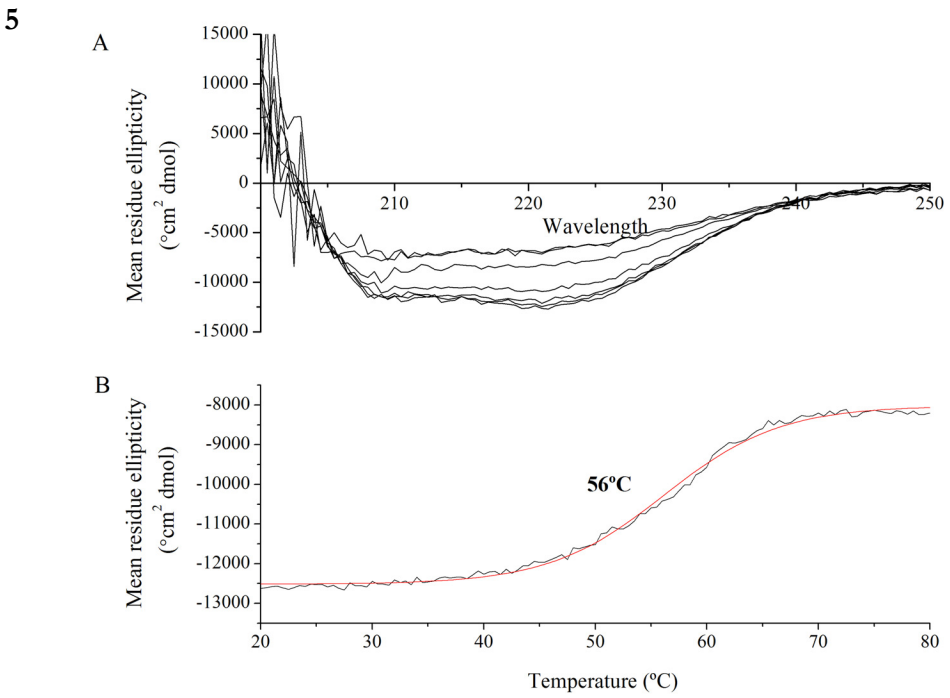
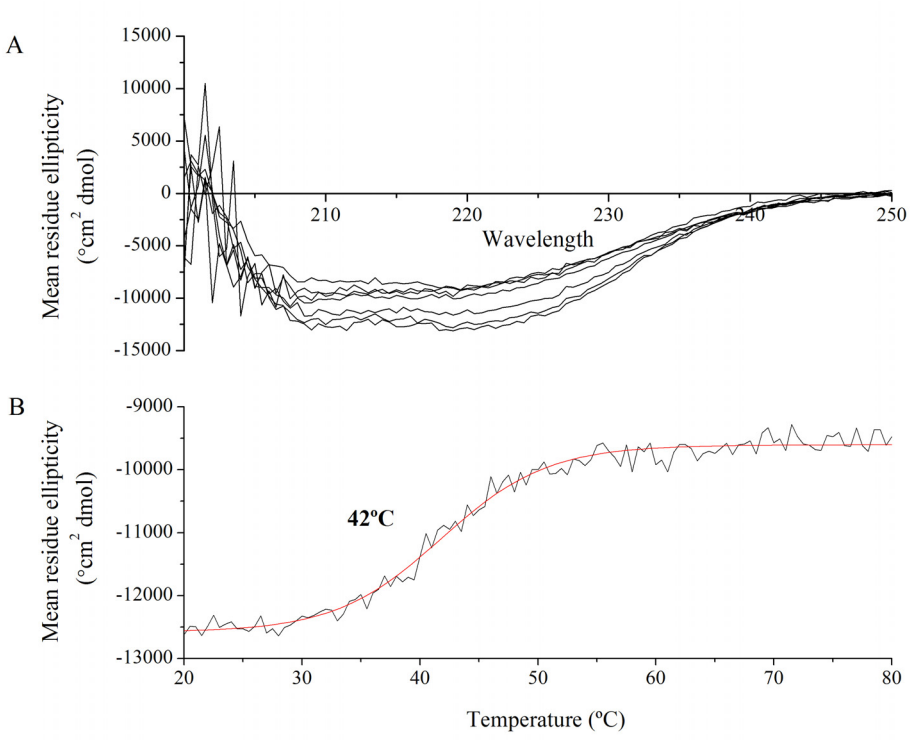
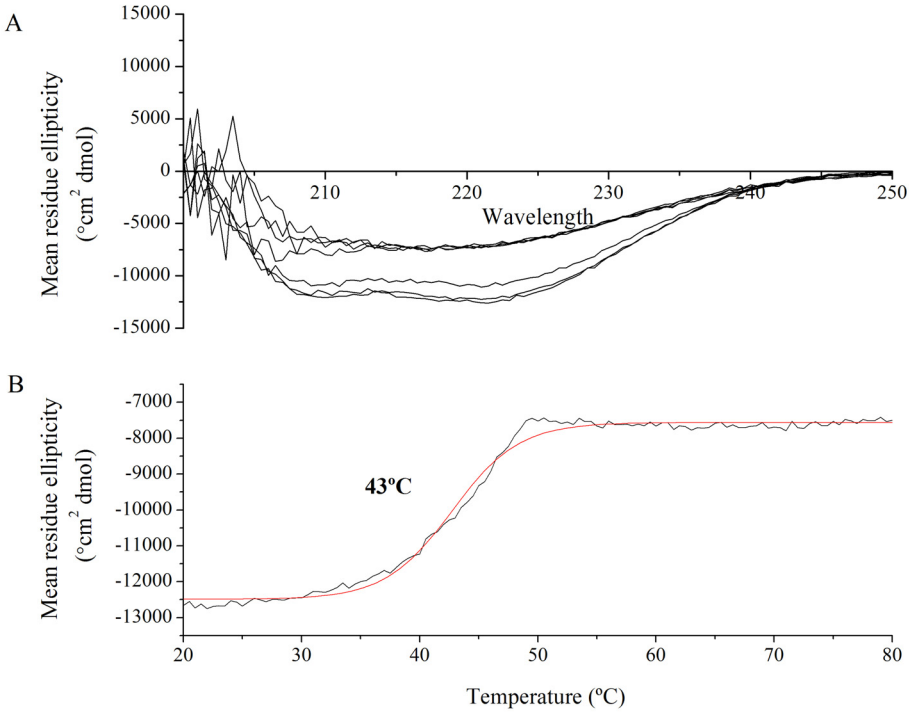


Figure A4.3 –Circular dichroism spectroscopy of **(1)** DHDPS Y107W **(2)** pET M11 DHDPS Y107W, **(3)** pET 151/D-TOPO DHDPS Y107W, **(4)** cleaved pET 151/D-TOPO DHDPS Y107W and **(5)** DHDPS Y107F. **A)** Wavelength scans plotting mean residue ellipticity vs. wavelength at 20°C, 30°C, 40°C, 50°C, 60°C, 70°C and 80°C. **B)** Mean residue ellipticity vs. temperature indicating melting of the protein. The red line corresponds to a sigmoid (Boltzmann) fit and the temperature shown is the midpoint of that fit.

1


A
Mean residue ellipticity
(°cm2 dmol)
Wavelength
B
42°C
Temperature (°C)

2


A
Mean residue ellipticity
(°cm2 dmol)
Wavelength
B
43°C
Temperature (°C)

3

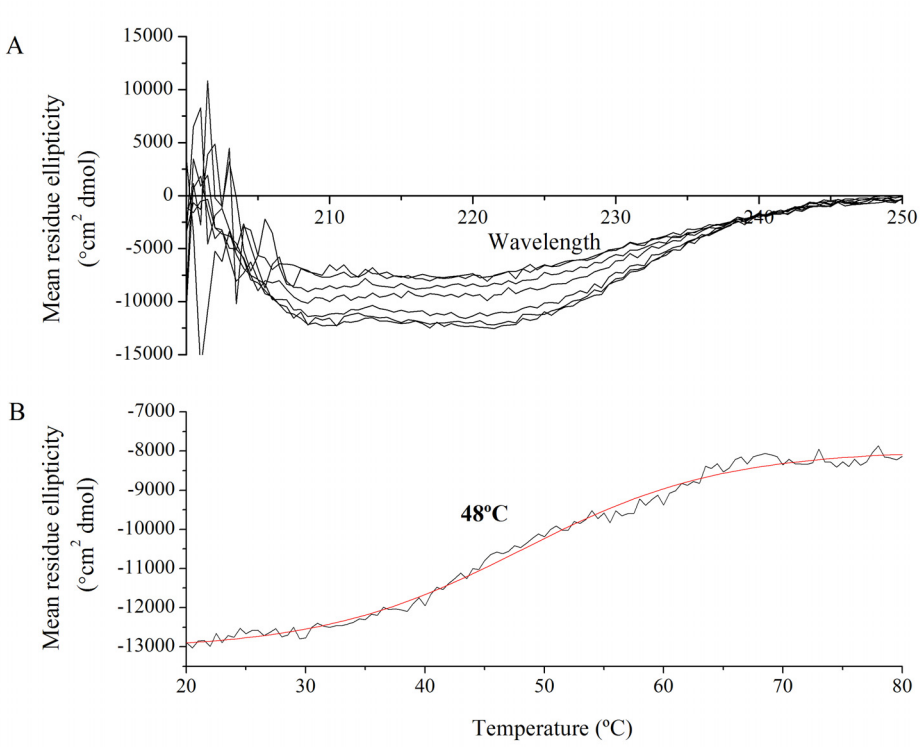


Figure A4.4 –Circular dichroism spectroscopy of **(1)** pET M11 DHDPS L197Y, **(2)** pET 151/D-TOPO DHDPS L197Y and **(3)** cleaved pET 151/D-TOPO DHDPS L197Y. **A)** Wavelength scans plotting mean residue ellipticity vs. wavelength at 20ºC, 30ºC, 40ºC, 50ºC, 60ºC, 70ºC and 80ºC. **B)** Mean residue ellipticity vs. temperature indicating melting of the protein. The red line corresponds to a sigmoid (Boltzmann) fit and the temperature shown is the midpoint of that fit.

A4.4 AMORPHOUS AGGREGATION

Variant	pH	Salt concentration (mM)	$Agg_{½}$ (minutes) ± SEM	Maximum absorbance (A.U.) ± SEM
DHDPS Y107W	2	0		0.07 ± 0.02
		20		0.06 ± 0.00
		40		0.10 ± 0.01
		100		0.12 ± 0.02
		500	4.86 ± 0.52	0.92 ± 0.08
	4.5	0	4.73 ± 0.32	1.14 ± 0.07
		20	6.09 ± 0.26	1.11 ± 0.03
		40	5.70 ± 0.28	1.21 ± 0.05
		100	4.81 ± 0.29	1.17 ± 0.04
		500	3.56 ± 0.33	1.11 ± 0.07
	7	0	15.86 ± 0.51	0.82 ± 0.03
		20	18.87 ± 0.23	0.83 ± 0.03

		40	17.93 ± 0.27	0.77 ± 0.03
		100	18.14 ± 0.35	0.80 ± 0.02
		500	17.47 ± 0.28	0.77 ± 0.03
	9.5	0	32.01 ± 2.12	0.42 ± 0.05
		20	25.98 ± 0.22	0.48 ± 0.02
		40	24.05 ± 0.22	0.59 ± 0.02
		100	23.05 ± 0.19	0.59 ± 0.06
		500	20.89 ± 0.17	0.67 ± 0.02
	11	0		0.04 ± 0.01
		20		0.09 ± 0.01
		40		0.10 ± 0.01
		100		0.07 ± 0.01
		500		0.38 ± 0.01
pET M11 DHDPS Y107W	2	0		0.10 ± 0.01
		20		0.08 ± 0.01
		40		0.12 ± 0.01
		100		0.15 ± 0.01
		500	15.62 ± 1.49	0.76 ± 0.06
	4.5	0	5.10 ± 0.73	1.16 ± 0.07
		20	5.70 ± 0.53	1.27 ± 0.09
		40	6.64 ± 0.51	1.18 ± 0.07
		100	6.48 ± 0.46	1.26 ± 0.06
		500	5.72 ± 0.23	1.14 ± 0.08
	7	0	13.77 ± 0.45	1.08 ± 0.04
		20	16.18 ± 0.10	1.04 ± 0.04
		40	16.59 ± 0.06	1.03 ± 0.04
		100	16.59 ± 0.14	1.04 ± 0.08
		500	17.71 ± 0.21	0.94 ± 0.04
	9.5	0	17.00 ± 0.41	0.96 ± 0.07
		20	19.01 ± 0.15	0.93 ± 0.05
		40	19.37 ± 0.08	1.04 ± 0.04
		100	18.51 ± 0.05	0.97 ± 0.05
		500	20.57 ± 0.22	0.91 ± 0.04
	11	0		0.14 ± 0.02
		20		0.13 ± 0.01
		40		0.15 ± 0.01
		100		0.14 ± 0.01
		500		0.50 ± 0.04
pET 151/D-TOPO DHDPS Y107W	2	0		0.06 ± 0.01
		20		0.09 ± 0.01
		40		0.09 ± 0.01
		100		0.14 ± 0.03
		500	12.98 ± 0.23	0.89 ± 0.05
	4.5	0	3.95 ± 0.43	1.47 ± 0.03
		20	4.98 ± 0.42	1.49 ± 0.03
		40	4.50 ± 0.59	1.44 ± 0.04
		100	6.33 ± 0.17	1.44 ± 0.03
		500	5.49 ± 0.27	1.22 ± 0.04
	7	0	14.52 ± 0.59	1.07 ± 0.02

		20	16.98 ± 0.29	1.01 ± 0.01
		40	17.16 ± 0.23	1.00 ± 0.02
		100	17.31 ± 0.19	1.03 ± 0.02
		500	18.76 ± 0.23	1.01 ± 0.02
	9.5	0	15.92 ± 0.59	0.86 ± 0.03
		20	20.59 ± 0.26	0.79 ± 0.01
		40	18.74 ± 0.22	0.88 ± 0.01
		100	19.02 ± 0.16	0.84 ± 0.02
		500	22.23 ± 0.24	0.86 ± 0.01
	11	0		0.08 ± 0.02
		20		0.07 ± 0.02
		40		0.10 ± 0.01
		100		0.06 ± 0.02
		500		0.67 ± 0.02
Cleaved pET 151/D-TOPO DHDPS Y107W	2	0		0.11 ± 0.02
		20		0.09 ± 0.02
		40		0.10 ± 0.01
		100		0.15 ± 0.02
		500	6.97 ± 1.52	1.04 ± 0.10
	4.5	0	2.78 ± 0.18	1.08 ± 0.07
		20	3.60 ± 0.15	1.11 ± 0.05
		40	3.69 ± 0.19	1.16 ± 0.05
		100	3.42 ± 0.26	1.07 ± 0.06
		500	3.30 ± 0.18	0.97 ± 0.05
	7	0	14.75 ± 0.60	1.07 ± 0.04
		20	16.85 ± 0.12	1.04 ± 0.02
		40	17.41 ± 0.06	1.04 ± 0.02
		100	17.31 ± 0.13	1.02 ± 0.02
		500	18.77 ± 0.24	0.97 ± 0.03
	9.5	0	16.69 ± 0.58	0.84 ± 0.04
		20	21.26 ± 0.09	0.78 ± 0.04
		40	19.12 ± 0.07	0.84 ± 0.02
		100	19.25 ± 0.10	0.84 ± 0.02
		500	20.53 ± 0.22	0.69 ± 0.03
	11	0		0.12 ± 0.01
		20		0.13 ± 0.01
		40		0.14 ± 0.01
		100		0.13 ± 0.01
		500		0.68 ± 0.03
DHDPS Y107F	2	0		0.10 ± 0.01
		20		0.09 ± 0.01
		40		0.12 ± 0.01
		100		0.15 ± 0.02
		500		0.79 ± 0.05
	4.5	0	23.00 ± 0.64	0.91 ± 0.02
		20	27.50 ± 0.28	0.90 ± 0.03
		40	28.29 ± 0.32	0.92 ± 0.03
		100	28.82 ± 0.31	0.88 ± 0.02
		500	26.09 ± 0.56	0.78 ± 0.03

	7	0	26.74 ± 0.50	0.64 ± 0.01
		20	30.42 ± 0.25	0.59 ± 0.01
		40	31.66 ± 0.25	0.55 ± 0.01
		100	32.55 ± 0.10	0.54 ± 0.01
		500	34.21 ± 1.93	0.52 ± 0.01
	9.5	0	36.24 ± 0.45	0.47 ± 0.03
		20	39.13 ± 0.40	0.42 ± 0.01
		40	38.31 ± 0.68	0.52 ± 0.02
		100	37.00 ± 0.50	0.45 ± 0.01
		500	31.70 ± 0.25	0.46 ± 0.00
	11	0		0.14 ± 0.01
		20		0.11 ± 0.01
		40		0.15 ± 0.02
		100		0.13 ± 0.01
		500		0.05 ± 0.03

Table A4.5 – The aggregation half life ($Agg_{½}$) and maxima of DHDPS Y107W pET M11 DHDPS Y107W, pET 151/D-TOPO DHDPS Y107W, cleaved pET 151/D-TOPO DHDPS Y107W and DHDPS Y107F as determined by the calculation of the midpoint of a sigmoid (Boltzmann) function fitted to the data as per figure 5.16. Errors were calculated as per the CD data (chapter 2, figure 2.17). The maximum absorbance was the highest reading over the initial 90 minutes of the assay.

Variant	pH	Salt concentration (mM)	$Agg_{½}$ (minutes) ± SEM	Maximum absorbance (AU) ± SEM
pET M11 DHDPS L197Y	2	0		0.09 ± 0.01
		20		0.10 ± 0.01
		40		0.11 ± 0.02
		100		0.14 ± 0.01
		500	11.98 ± 0.86	0.78 ± 0.05
	4.5	0	2.46 ± 0.14	1.22 ± 0.07
		20	2.82 ± 0.08	1.22 ± 0.03
		40	3.69 ± 0.08	1.34 ± 0.02
		100	4.11 ± 0.03	1.18 ± 0.10
		500	5.06 ± 0.18	1.16 ± 0.04
	7	0	10.97 ± 0.49	1.10 ± 0.01
		20	13.24 ± 0.08	1.10 ± 0.02
		40	13.42 ± 0.09	1.07 ± 0.01
		100	13.69 ± 0.09	1.09 ± 0.01
		500	16.80 ± 0.18	1.03 ± 0.01
	9.5	0	14.28 ± 0.52	1.05 ± 0.03
		20	16.11 ± 0.05	0.93 ± 0.01
		40	16.38 ± 0.22	1.10 ± 0.05
		100	16.04 ± 0.07	0.93 ± 0.02
		500	19.15 ± 0.22	0.91 ± 0.01
	11	0		0.14 ± 0.01
		20		0.11 ± 0.01

		40		0.14 ± 0.01
		100		0.10 ± 0.01
		500		0.48 ± 0.02
pET 151/D-TOPO DHDPS L197Y	2	0		0.14 ± 0.03
		20		0.13 ± 0.01
		40		0.16 ± 0.01
		100		0.19 ± 0.02
		500	11.63 ± 0.27	1.10 ± 0.06
	4.5	0	2.17 ± 0.07	1.41 ± 0.03
		20	2.82 ± 0.10	1.42 ± 0.03
		40	3.01 ± 0.14	1.47 ± 0.01
		100	3.71 ± 0.17	1.47 ± 0.04
		500	5.61 ± 0.20	1.37 ± 0.05
	7	0	11.02 ± 0.51	1.36 ± 0.03
		20	13.12 ± 0.27	1.32 ± 0.02
		40	13.70 ± 0.23	1.30 ± 0.01
		100	14.16 ± 0.22	1.32 ± 0.05
		500	17.46 ± 0.41	1.27 ± 0.05
	9.5	0	12.57 ± 0.61	1.06 ± 0.03
		20	14.76 ± 0.26	1.04 ± 0.03
		40	15.22 ± 0.20	1.06 ± 0.03
		100	15.70 ± 0.25	1.05 ± 0.01
		500	19.69 ± 0.34	1.04 ± 0.01
	11	0		0.12 ± 0.01
		20		0.08 ± 0.01
		40		0.10 ± 0.01
		100		0.10 ± 0.00
		500		0.64 ± 0.01
Cleaved pET 151/D-TOPO DHDPS L197Y	2	0		0.08 ± 0.01
		20		0.09 ± 0.00
		40		0.08 ± 0.01
		100		0.15 ± 0.02
		500		0.86 ± 0.09
	4.5	0	2.04 ± 0.18	1.27 ± 0.05
		20	2.60 ± 0.25	1.22 ± 0.03
		40	2.86 ± 0.22	1.26 ± 0.02
		100	3.49 ± 0.28	1.31 ± 0.05
		500	4.55 ± 0.28	1.23 ± 0.04
	7	0	12.20 ± 0.62	1.14 ± 0.04
		20	14.18 ± 0.33	1.10 ± 0.03
		40	14.59 ± 0.29	1.11 ± 0.01
		100	15.31 ± 0.32	1.13 ± 0.01
		500	18.27 ± 0.36	1.02 ± 0.02
	9.5	0	14.02 ± 0.63	0.96 ± 0.03
		20	16.66 ± 0.33	0.81 ± 0.02
		40	16.60 ± 0.24	0.89 ± 0.03
		100	17.03 ± 0.30	0.87 ± 0.01
		500	21.91 ± 0.40	0.83 ± 0.01
	11	0		0.07 ± 0.01

		20		0.06 ± 0.01
		40		0.11 ± 0.02
		100		0.09 ± 0.01
		500		0.63 ± 0.02

Table A4.6 – The aggregation half life ($Agg_{½}$) and maxima of pET M11 DHDPS L197Y, pET 151/D-TOPO DHDPS L197Y and cleaved pET 151/D-TOPO DHDPS L197Y as determined by the calculation of the midpoint of a sigmoid (Boltzmann) function fitted to the data as per figure 5.18. Errors were calculated as per the CD data (chapter 2, figure 2.17). The maximum absorbance was the highest reading over the initial 90 minutes of the assay.

	Degrees of freedom	Sum of squares	Mean squares	F value	Pr (>F)
Protein	8	69751	8719	4957	$< 2.2 \times 10^{-16}$
Salt	4	160	40	23	$< 2.2 \times 10^{-16}$
pH	1	2669	2669	1518	$< 2.2 \times 10^{-16}$
Protein:salt	32	4864	152	86	$< 2.2 \times 10^{-16}$
Protein:pH	8	1629	204	116	$< 2.2 \times 10^{-16}$
salt:pH	4	347	87	49	$< 2.2 \times 10^{-16}$
Protein:pH:salt	32	935	29	17	$< 2.2 \times 10^{-16}$
Residuals	450	791	2		

Table A4.7 - Analysis of deviance table for the statistical analysis of the $Agg_{½}$ of DHDPS Y107W, L197Y, their polyhistidine tagged variants and DHDPS Y107F.

	Degrees of freedom	Sum of squares	Mean squares	F value	Pr (>F)
Protein	8	14.44	1.81	17.66	$< 2.2 \times 10^{-16}$
Salt	4	31.21	7.8	76.35	$< 2.2 \times 10^{-16}$
pH	4	194.6	48.66	476.19	$< 2.2 \times 10^{-16}$
Protein:salt	32	0.73	0.02	0.23	1
Protein:pH	32	10.78	0.34	3.3	2.9×10^{-9}
salt:pH	16	28.41	1.78	17.38	1
Protein:pH:salt	128	2.99	0.02	0.23	$< 2.2 \times 10^{-16}$
Residuals	1125	114.96	0.1		

Table A4.8 - Analysis of deviance table for the statistical analysis of the maximum absorbance of DHDPS Y107W, L197Y, their polyhistidine tagged variants and DHDPS Y107F.

A4.5 β-SHEET-SPECIFIC AGGREGATION

Variant	pH	Salt concentration (mM)	$Agg_{½}$ (minutes) ± SEM	Maximum fluorescence (RFU) ± SEM
DHDPS Y107W	2	0		-120 ± 16
		40		-116 ± 26
		500		274 ± 35
	4.5	0		1139 ± 49
		40		1175 ± 86
		500		2176 ± 87
	7	0	47.3 ± 6.6	3908 ± 264
		40	62.7 ± 2.0	3833 ± 197
		500	67.3 ± 1.0	2794 ± 138
	9.5	0	48.0 ± 3.5	988 ± 16
		40	54.3 ± 1.8	1099 ± 42
		500	75.0 ± 2.4	1412 ± 150
	11	0	20.6 ± 2.6	7 ± 9
		40	35.9 ± 2.6	13 ± 7
		500	27.4 ± 13.4	552 ± 54
pET M11 DHDPS Y107W	2	0		-35 ± 16
		40		-30 ± 20
		500		480 ± 46
	4.5	0	31.2 ± 8.1	1505 ± 287
		40	35.9 ± 4.9	1632 ± 195
		500	68.8 ± 2.2	2673 ± 309
	7	0	61.7 ± 4.0	6726 ± 680
		40	63.4 ± 1.3	6612 ± 478
		500	75.7 ± 3.1	5582 ± 411
	9.5	0	13.1 ± 2.4	1490 ± 84
		40	17.7 ± 4.8	1630 ± 58
		500	19.2 ± 1.8	2249 ± 163
	11	0		-64 ± 11
		40		-43 ± 18
		500		563 ± 106
pET 151/D-TOPO DHDPS Y107W	2	0		-35 ± 25
		40		21 ± 28
		500		816 ± 29
	4.5	0	42.1 ± 6.6	3465 ± 279
		40	54.3 ± 1.8	4489 ± 132
		500	77.5 ± 1.7	6563 ± 264
	7	0	49.4 ± 2.9	11165 ± 767
		40	54.8 ± 2.3	11511 ± 805
		500	59.4 ± 2.7	8949 ± 503
	9.5	0	23.3 ± 1.9	1477 ± 87
		40	26.2 ± 1.0	1722 ± 101
		500	24.9 ± 1.0	2517 ± 60
	11	0		-97 ± 20
		40		-59 ± 22

		500		152 ± 33
Cleaved pET 151/D-TOPO DHDPS Y107W	2	0		11 ± 66
		40		57 ± 60
		500		923 ± 237
	4.5	0	62.2 ± 1.7	3470 ± 844
		40	68.5 ± 1.2	4231 ± 783
		500	80.5 ± 1.5	6892 ± 1462
	7	0	48.1 ± 1.7	10267 ± 1508
		40	45.5 ± 0.9	7708 ± 1124
		500	35.8 ± 8.3	5617 ± 545
	9.5	0	33.4 ± 0.3	2649 ± 353
		40	33.6 ± 0.1	2641 ± 274
		500	28.5 ± 1.4	2219 ± 246
	11	0		-30 ± 24
		40		-89 ± 22
		500		958 ± 113
DHDPS Y107F	2	0		-26 ± 17
		40		-23 ± 13
		500		-12 ± 35
	4.5	0	67.4 ± 2.2	830 ± 134
		40	64.7 ± 1.2	1017 ± 166
		500	61.7 ± 2.0	2372 ± 248
	7	0	61.5 ± 0.8	4742 ± 433
		40	66.5 ± 2.3	3863 ± 289
		500	70.3 ± 1.2	2532 ± 167
	9.5	0	50.6 ± 3.5	1805 ± 706
		40	52.1 ± 2.9	1466 ± 555
		500	52.0 ± 1.5	1154 ± 223
	11	0		-15 ± 3
		40		-14 ± 16
		500		-21 ± 5

Table A4.9 – The speed of β-sheet-specific aggregation ($\beta\text{-agg}_{½}$) of DHDPS Y107W, pET M11 DHDPS Y107W, pET 151/D-TOPO DHDPS Y107W, cleaved pET 151/D-TOPO DHDPS Y107W and DHDPS Y107F as determined by the calculation of the $t_{½}$ of a sigmoid (Boltzmann) function fitted to the ThT fluorescence data (chapter 5, figure 5.20). Errors were calculated as per the CD data (chapter 2, figure 2.17). The maximum absorbance was the highest reading over the initial 300 minutes of the assay.

Variant	pH	Salt concentration (mM)	$Agg_{½}$ (minutes) ± SEM	Maximum fluorescence (RFU) ± SEM
pET M11 DHDPS L197Y	2	0		-47 ± 32
		40		-102 ± 23
		500		81 ± 16
	4.5	0	53.0 ± 4.7	866 ± 61
		40	50.2 ± 5.4	866 ± 115
		500	81.9 ± 3.2	1395 ± 98
	7	0	66.5 ± 4.2	2336 ± 228
		40	64.9 ± 3.0	2324 ± 102
		500	69.7 ± 6.1	2057 ± 79
	9.5	0	17.1 ± 0.0	877 ± 71
		40	2.3 ± 1.9	890 ± 30
		500	74.3 ± 3.0	1024 ± 67
	11	0		-38 ± 13
		40		-2 ± 15
		500		275 ± 13
pET 151/D-TOPO DHDPS L197Y	2	0		165 ± 35
		40		233 ± 17
		500		1502 ± 327
	4.5	0	53.9 ± 1.3	3697 ± 380
		40	55.1 ± 1.7	4433 ± 435
		500	62.2 ± 1.1	7554 ± 757
	7	0	41.3 ± 1.8	11105 ± 655
		40	32.5 ± 5.1	9086 ± 408
		500	41.9 ± 4.7	7815 ± 715
	9.5	0	19.1 ± 0.4	3134 ± 286
		40	24.6 ± 1.2	2960 ± 239
		500	43.7 ± 0.6	3665 ± 386
	11	0		4 ± 18
		40		-11 ± 17
		500		301 ± 74
Cleaved pET 151/D-TOPO DHDPS L197Y	2	0		-111 ± 18
		40		-88 ± 12
		500		527 ± 33
	4.5	0	52.1 ± 2.5	3280 ± 171
		40	57.5 ± 2.2	3462 ± 121
		500	66.1 ± 6.1	5264 ± 181
	7	0	58.2 ± 2.2	7356 ± 643
		40	57.8 ± 3.4	6752 ± 565
		500	55.7 ± 7.0	5646 ± 342
	9.5	0	35.1 ± 0.8	2491 ± 330
		40	39.0 ± 1.5	2587 ± 348
		500	37.6 ± 1.5	2934 ± 232
	11	0		-46 ± 13
		40		-44 ± 11
		500		4 ± 31

Table A4.10 – The speed of β-sheet-specific aggregation (β-$agg_{½}$) of pET M11 DHDPS L197Y, pET 151/D-TOPO DHDPS L197Y and cleaved pET 151/D-TOPO DHDPS L197Y as determined by the calculation of the $t_{½}$ of a sigmoid (Boltzmann) function fitted to the ThT fluorescence data (chapter 5, figure 5.21). Errors were calculated as per the CD data (chapter 2, figure 2.17). The maximum absorbance was the highest reading over the initial 300 minutes of the assay.

	Degrees of freedom	Sum of squares	Mean squares	F value	Pr (>F)
Protein	8	20629	2579	36	$< 2.2 \times 10^{-16}$
pH	1	63	63	0.9	3.4×10^{-1}
Salt	2	11868	5934	83	$< 2.2 \times 10^{-16}$
Protein:pH	8	18278	2285	32	$< 2.2 \times 10^{-16}$
Protein:salt	16	5686	355	5	4.7×10^{-9}
pH:salt	2	1941	971	14	2.2×10^{-6}
Protein:pH:salt	16	4724	295	4	3.6×10^{-7}
Residuals	270	19189	71		

Table A4.11 - Analysis of deviance table for the statistical analysis of the β-$agg_{½}$ of DHDPS Y107W, L197Y, their polyhistidine tagged variants and Y107F. pH and salt were included in the model because they had an effect on the fluorescence however this was not relevant due to their affect on ThT binding. The main effect of protein was the independent effect after controlling between the differences across pH and salt concentrations.

	Degrees of freedom	Sum of squares	Mean squares	F value	Pr (>F)
Protein	8	720925949	102989421	187	$< 2.2 \times 10^{-16}$
pH	4	3303723852	824930963	1503	$< 2.2 \times 10^{-16}$
Salt	2	25804200	12902100	23	1.5×10^{-10}
Protein:pH	28	794613286	28379046	52	$< 2.2 \times 10^{-16}$
Protein:salt	14	43065367	3076098	6	3.0×10^{-10}
pH:salt	8	45327364	5665920	10	1.4×10^{-13}
Protein:pH:salt	56	62719131	1119984	2	2.8×10^{-5}
Residuals	600	329565120	549275		

Table A3.14 - Analysis of deviance table for the statistical analysis of the maximum fluorescence of DHDPS Y107W, L197Y, their polyhistidine tagged variants and Y107F. As discussed above, pH and salt were included in the model because they had an effect on the fluorescence however this was not relevant due to their affect on ThT binding.

APPENDIX 5

SUPPLEMENTARY INFORMATION FOR CHAPTER 6

A5.1 ALGORITHMIC PREDICTION

	Wild-type DHDPS	pET 151/D-TOPO DHDPS Q90L/Y107W	cleaved pET 151/D-TOPO DHDPS Q90L/Y107W	pET 151/D-TOPO DHDPS Q90L	pET 151/D-TOPO DHDPS Y107W
Overall hydrophobicity at pH 7	95.38	95.25	91.33	95.91	98.37
Overall charge at pH 7	-5	-7	-6	-7	-7
Overall β-sheet propensity	1090.06	1195.56	1105.85	1195.59	1195.37
Overall α-helical propensity	1133.27	1250.4	1151.41	1250.53	1250.15
Number of hydrophobic/hydrophilic patterns	2	6	6	6	2
Absolute amyloid aggregation rate	-4.07	-2.66	-2.54	-2.67	-4.26
Intrinsic aggregation propensity	-6.12	-5.25	-4.67	-5.26	-6.86
Average ± SD	-5.46 ± 1.45	-5.81 ± 1.53	-5.52 ± 1.47	-5.81 ± 1.53	-5.81 ± 1.53
Zagg score	-0.46	0.370186	0.574769	0.359207	-0.68411

Table A5.1 – The results of the Zyggregator algorithm for wild-type DHDPS, pET 151/D-TOPO DHDPS Q90L/Y107W, cleaved pET 151/D-TOPO DHDPS Q90L/Y107W, pET 151/D-TOPO DHDPS Q90L and pET 151/D-TOPO DHDPS Y107W.

A5.2 DIFFERENTIAL SCANNING FLUORIMETRY

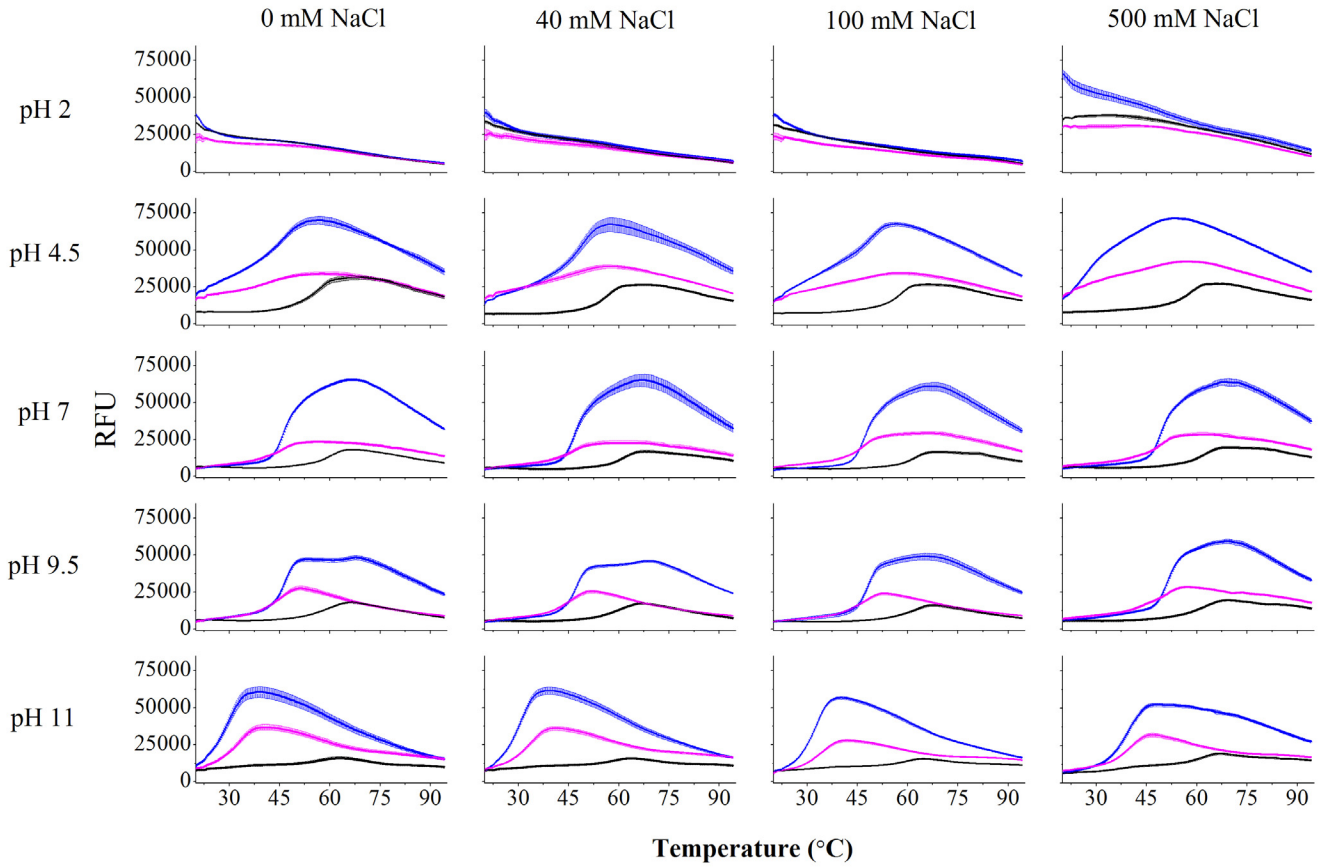


Figure A5.1 – The thermal denaturation of wild-type *E. coli* DHDPS (black), pET 151/D-TOPO polyhistidine-tagged *E. coli* DHDPS Q90L/Y107W (blue) and cleaved pET 151/D-TOPO polyhistidine-tagged *E. coli* DHDPS Q90L/Y107W (pink) in 100 mM phosphate buffer at pH 2, pH 4.5, pH 7, pH 9.5 and pH 11 containing 0 mM NaCl, 40 mM NaCl, 100 mM NaCl and 500 mM NaCl as monitored by SYPRO Orange fluorescence. The data plotted are the increase in fluorescence, as monitored by the BioRad IQ5.

Variant	pH	Salt concentration (mM)	Mean T_m 1 (°C)	Mean T_m 2 (°C)
pET 151/D-TOPO DHDPS Q90L/Y107W	4.5	0	47.6 ± 0.0	
		40	47.6 ± 0.0	
		100	47.6 ± 0.0	
		500	26.5 ± 0.1	
	7	0	46.0 ± 0.1	56.0 ± 0.0
		40	46.3 ± 0.1	56.0 ± 0.0
		100	46.4 ± 0.2	55.4 ± 0.3
		500	47.2 ± 0.3	55.7 ± 0.3
	9.5	0	46.4 ± 0.5	65.6 ± 0.3
		40	47.4 ± 0.1	65.3 ± 0.3
		100	47.5 ± 0.1	62.9 ± 0.0
		500	47.6 ± 0.0	55.4 ± 0.3
	11	0	29.6 ± 0.0	
		40	30.8 ± 0.0	
		100	30.8 ± 0.0	
		500	30.8 ± 0.0	
Cleaved pET 151/D-TOPO DHDPS Q90L/Y107W	4.5	0	40.3 ± 1.0	
		40	39.3 ± 1.0	
		100	40.3 ± 1.0	
		500	27.2 ± 0.0	
	7	0	46.4 ± 0.2	
		40	46.4 ± 0.0	
		100	46.6 ± 0.1	
		500	47.1 ± 0.1	
	9.5	0	45.4 ± 0.1	
		40	45.2 ± 0.0	
		100	45.3 ± 0.1	
		500	45.1 ± 0.1	
	11	0	33.1 ± 0.3	
		40	32.8 ± 0.2	
		100	33.0 ± 0.2	
		500	33.5 ± 0.6	

Table A5.2 – The thermal denaturation temperatures of pET 151/D-TOPO DHDPS Q90L/Y107W and cleaved pET 151/D-TOPO DHDPS Q90L/Y107W in 100 mM phosphate buffer as monitored by SYPRO Orange fluorescence and as assessed from the derivation of the fluorescence data.

A5.3 CIRCULAR DICHROISM SPECTROSCOPY

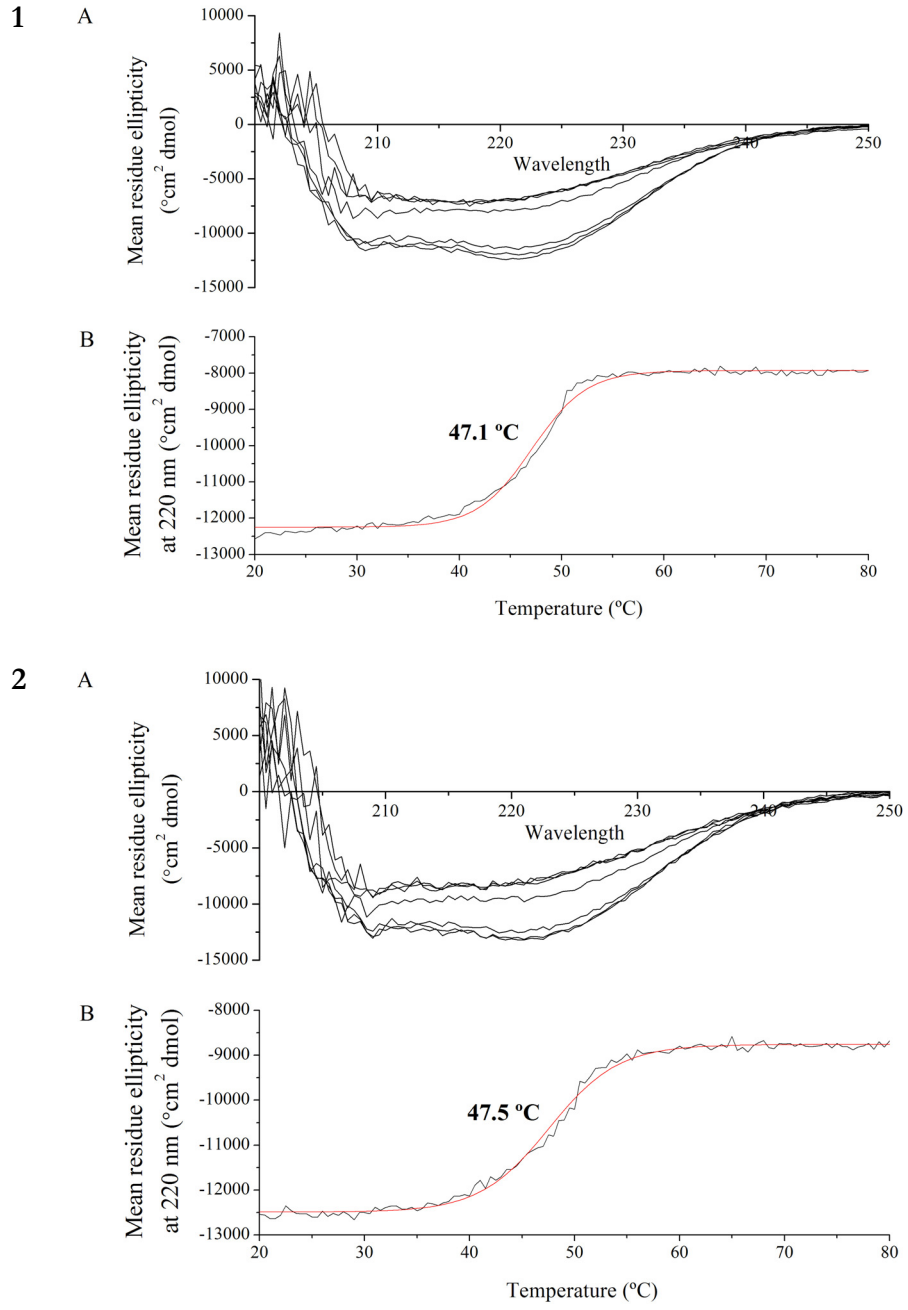


Figure A5.2 – Circular dichroism spectroscopy of **(1)** pET 151/D-TOPO DHDPS Q90L/Y107W and **(2)** cleaved pET 151/D-TOPO DHDPS Q90L/Y107W. **A)** Wavelength scans plotting mean residue ellipticity vs. wavelength at 20ºC, 30ºC, 40ºC, 50ºC, 60ºC, 70ºC and 80ºC. **B)** Mean residue ellipticity vs. temperature indicating melting of the protein. The red line corresponds to a sigmoid (Boltzmann) fit and the temperature shown is the midpoint of that fit.

A5.4 AMORPHOUS AGGREGATION

Variant	pH	Salt concentration (mM)	$Agg_{½}$ (minutes) ± SEM	Maximum absorbance (AU) ± SEM
pET 151/D-TOPO DHDPS Q90L/Y107W	2	0		0.10 ± 0.01
		20		0.11 ± 0.01
		40		0.12 ± 0.02
		100		0.16 ± 0.02
		500	14.65 ± 1.01	0.84 ± 0.05
	4.5	0	4.47 ± 0.44	1.36 ± 0.05
		20	31.97 ± 0.39	1.31 ± 0.08
		40	5.94 ± 0.67	1.43 ± 0.06
		100	5.28 ± 0.42	1.47 ± 0.06
		500	5.94 ± 0.11	1.29 ± 0.06
	7	0	14.61 ± 0.50	1.06 ± 0.04
		20	17.13 ± 0.17	1.01 ± 0.03
		40	17.50 ± 0.16	1.06 ± 0.04
		100	17.48 ± 0.13	1.03 ± 0.02
		500	19.30 ± 0.23	0.99 ± 0.02
	9.5	0	16.94 ± 0.82	0.84 ± 0.09
		20	19.42 ± 0.21	0.79 ± 0.04
		40	19.06 ± 0.18	0.87 ± 0.03
		100	18.77 ± 0.16	0.88 ± 0.02
		500	20.75 ± 0.17	0.90 ± 0.02
	11	0		0.05 ± 0.01
		20		0.03 ± 0.01
		40		0.06 ± 0.01
		100		0.07 ± 0.02
		500		0.86 ± 0.02
Cleaved pET 151/D-TOPO DHDPS Q90L/Y107W	2	0		0.12 ± 0.01
		20		
		40		0.11 ± 0.01
		100		
		500	7.82 ± 0.61	0.72 ± 0.04
	4.5	0	2.52 ± 0.65	0.81 ± 0.03
		20		
		40	3.28 ± 0.49	0.89 ± 0.05
		100		
		500	2.71 ± 0.60	0.81 ± 0.04
	7	0	14.44 ± 0.22	0.69 ± 0.03
		20		
		40	16.01 ± 0.23	0.68 ± 0.05
		100		
		500	15.96 ± 0.26	0.68 ± 0.04
	9.5	0	40.46 ± 0.33	0.19 ± 0.03
		20		
		40	21.93 ± 0.27	0.45 ± 0.04

		100		
		500	18.15 ± 0.65	0.49 ± 0.05
	11	0		0.08 ± 0.01
		20		
		40		0.10 ± 0.00
		100		
		500		0.30 ± 0.03

Table A5.4 – The aggregation half life ($Agg_{½}$) and maxima of pET 151/D-TOPO DHDPS Q90L/Y107W and cleaved pET 151/D-TOPO DHDPS Q90L/Y107W determined by the calculation of the $t_{½}$ of a sigmoid function fitted to the light scattering data (chapter 6, figure 6.6). Errors were calculated as per the CD data (figure 3.16). The maximum absorbance was the maximum reached over the initial 90 minutes of the assay. The errors were calculated from the mean of the six replicates.

	Degrees of freedom	Sum of squares	Mean squares	F value	Pr (>F)
Protein	2	18019	9009	4150	$< 2.2 \times 10^{-16}$
Salt	2	903	451	208	$< 2.2 \times 10^{-16}$
pH	1	750	750	346	$< 2.2 \times 10^{-16}$
Protein:salt	4	2599	650	299	$< 2.2 \times 10^{-16}$
Protein:pH	2	778	389	179	$< 2.2 \times 10^{-16}$
salt:pH	2	358	179	82	$< 2.2 \times 10^{-16}$
Protein:pH:salt	4	332	83	38	$< 2.2 \times 10^{-16}$
Residuals	90	195	2		

Table A5.5 - Analysis of deviance table for the statistical analysis of the $Agg_{½}$ of the variants of DHDPS Q90L/Y107W.

	Degrees of freedom	Sum of squares	Mean squares	F value	Pr (>F)
Protein	2	2.96	1.48	15.92	2.54×10^{-7}
Salt	4	7.1	1.78	19.1	4.0×10^{-14}
pH	4	50.19	12.55	134.96	$< 2.2 \times 10^{-16}$
Protein:salt	6	0.62	0.104	1.12	0.35
Protein:pH	8	2.41	0.3	3.24	1.5×10^{-3}
salt:pH	16	5.07	0.32	3.41	1.3×10^{-5}
Protein:pH:salt	128	2.99	0.02	0.23	$< 2.2 \times 10^{-16}$
Residuals	325	30.21	0.09		

Table A2.6 - Analysis of deviance table for the statistical analysis of the maximum absorbance of the variants of DHDPS Q90L/Y107W.

A5.5 β-SHEET-SPECIFIC AGGREGATION

Variant	pH	Salt concentration (mM)	β-$agg_{½}$ (minutes) ± SEM	Maximum fluorescence (RFU) ± SEM
pET 151/D-TOPO DHDPS Q90L/Y107W	2	0		-153 ± 41
		40		-186 ± 13
		500		373 ± 20
	4.5	0	40.9 ± 1.9	1192 ± 176
		40	50.2 ± 1.2	1713 ± 60
		500	72.6 ± 0.8	2997 ± 382
	7	0	53.0 ± 1.3	8023 ± 155
		40	57.0 ± 1.1	7956 ± 217
		500	62.0 ± 2.2	6153 ± 330
	9.5	0	32.1 ± 1.2	4195 ± 433
		40	32.9 ± 0.2	3818 ± 260
		500	38.5 ± 0.7	3273 ± 121
	11	0		-54 ± 17
		40		26 ± 15
		500		825 ± 54
Cleaved pET 151/D-TOPO DHDPS Q90L/Y107W	2	0		409 ± 229
		40		-28 ± 26
		500		230 ± 35
	4.5	0		304 ± 42
		40	73.7 ± 4.5	647 ± 83
		500	167.7 ± 2.5	1819 ± 122
	7	0	44.6 ± 9.8	3004 ± 293
		40	48.6 ± 9.3	3110 ± 281
		500	81.8 ± 4.8	2500 ± 119
	9.5	0	20.5 ± 0.4	1008 ± 76
		40	24.3 ± 1.0	1026 ± 50
		500	39.2 ± 4.9	1210 ± 68
	11	0		-52 ± 10
		40		-39 ± 11
		500		408 ± 74

Table A5.7 –The β-sheet-specific aggregation half life (β-$agg_{½}$) of pET 151/D-TOPO DHDPS Q90L/Y107W and cleaved pET 151/D-TOPO DHDPS Q90L/Y107W as determined by the calculation of the $t_{½}$ of a sigmoid (Boltzmann) function fitted to the ThT fluorescence data (chapter 6, figure 6.8). Errors were calculated as per the CD data (chapter 3, figure 3.16). The maximum fluorescence was the maximum reached over the initial 300 minutes of the assay. The errors were calculated from the mean of the six replicates.

	Degrees of freedom	Sum of squares	Mean squares	F value	Pr (>F)
Protein	2	14051	7026	71	$< 2.2 \times 10^{-16}$
pH	1	1981	1981	20	2.2×10^{-5}
Salt	2	27348	13674	138	$< 2.2 \times 10^{-16}$
Protein:pH	2	20691	10345	105	$< 2.2 \times 10^{-16}$
Protein:salt	4	11635	2909	29	1.2×10^{-15}
pH:salt	2	4172	2086	21	3.0×10^{-8}
Protein:pH:salt	4	3763	941	9	1.8×10^{-6}
Residuals	90	8891	99		

Table A2.8 - Analysis of deviance table for the statistical analysis of the β-agg$_{½}$ of the variants of DHDPS Q90L/Y107W. pH and salt were included in the model because they had an effect on the fluorescence; however, this was not relevant due to their affect on ThT binding. The main effect of protein was the independent effect after controlling between the differences across pH and salt concentrations.

	Degrees of freedom	Sum of squares	Mean squares	F value	Pr (>F)
Protein	2	144210864	72105432	115	$< 2.2 \times 10^{-16}$
pH	4	1127830409	281957602	450	$< 2.2 \times 10^{-16}$
Salt	4	36946856	9236714	15	6.3×10^{-11}
Protein:pH	8	199648780	24956097	40	$< 2.2 \times 10^{-16}$
Protein:salt	4	10523266	2630817	4	2.6×10^{-3}
pH:salt	16	52533424	3283339	5	1.3×10^{-9}
Protein:pH:salt	16	15399996	962500	2	8.7×10^{-2}
Residuals	275	172281154	626477		

Table A2.9 - Analysis of deviance table for the statistical analysis of the maximum fluorescence of the variants of DHDPS Q90L/Y107W. As discussed above, pH and salt were included in the model because they had an effect on the fluorescence; however, this was not relevant due to their affect on ThT binding.